

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 06:30:32 ; Search time 16.3133 Seconds  
(without alignments)  
1028.511 Million cell updates/sec

Title: US-09-980-464-11\_COPY\_57\_309  
Perfect score: 1342  
Sequence: 1 YEFLETGKGTGKVKKARE.....VNPTTRATLEVDVASHWVWV 253

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1342	100.0	630	4	US-10-355-975A-38
2	1342	100.0	631	4	US-09-579-664B-11
3	1342	100.0	631	4	US-10-355-975A-11
4	649.5	48.4	729	2	US-08-677-298-2
5	649.5	48.4	729	4	US-09-523-849-33
6	647.5	48.2	722	4	US-09-984-890-4
7	647.5	48.2	722	4	US-10-274-194-4
8	647.5	48.2	745	4	US-09-523-849-36
9	646.5	48.2	713	4	US-09-538-092-1022
10	646.5	48.2	724	4	US-09-984-890-2
11	646.5	48.2	724	4	US-10-274-194-2
12	644	48.0	512	4	US-09-633-328B-2
13	642.5	47.9	722	4	US-08-817-832B-32
14	642.5	47.9	793	4	US-09-523-849-32
15	631	47.0	511	4	US-09-633-328B-4
16	629	46.9	779	4	US-08-817-832B-31
17	625	46.6	257	3	US-09-101-146-6
18	619.5	46.2	257	2	US-07-857-224B-25
19	619.5	46.2	633	3	US-08-557-006C-43
20	619.5	46.2	633	4	US-09-538-092-212
21	619.5	46.2	633	4	US-09-633-328B-3
22	619.5	46.2	633	4	US-09-824-735-3
23	618.5	46.1	345	3	US-08-101-146-1
24	615.5	45.9	1203	4	US-09-799-875-5
25	607.5	45.3	604	4	US-09-523-849-35
26	604	45.0	552	4	US-09-824-735-4
27	601	44.8	552	3	US-08-557-006C-40

28	600	44.7	552	4	US-09-538-092-1212	Sequence 1212, Ap
29	597.5	44.5	418	4	US-09-248-796A-18441	Sequence 18441, A
30	597.5	44.5	776	4	US-09-523-849-34	Sequence 34, Appl
31	562.5	41.9	778	4	US-10-116-326-2	Sequence 2, Appl
32	562.5	41.9	778	4	US-10-003-690-2	Sequence 2, Appl
33	556.5	41.5	668	4	US-09-930-181-2	Sequence 2, Appl
34	523	39.0	1064	4	US-09-538-092-154	Sequence 154, App
35	522	38.9	260	2	US-07-857-224B-28	Sequence 28, Appl
36	515.5	38.4	160	4	US-09-270-767-32886	Sequence 32886, A
37	513	38.2	260	2	US-07-857-224B-27	Sequence 27, Appl
38	507.5	37.8	703	4	US-10-116-326-6	Sequence 6, Appl
39	504.5	37.6	436	4	US-09-734-673-2	Sequence 2, Appl
40	504.5	37.6	436	4	US-09-523-849-2	Sequence 2, Appl
41	502.5	37.4	446	4	US-09-824-735-2	Sequence 2, Appl
42	499	37.2	391	4	US-09-248-796A-18435	Sequence 18435, A
43	498	37.1	915	4	US-09-538-092-63	Sequence 63, Appl
44	493	36.7	360	4	US-09-248-796A-20497	Sequence 20497, A
45	485.5	36.2	603	4	US-09-930-181-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1  
US-10-355-975A-38  
; Sequence 38, Application US/10355975A  
; Patent No. 6759223  
; GENERAL INFORMATION:  
; APPLICANT: Immunex Corporation  
; APPLICANT: Bird, Timothy A.  
; APPLICANT: Virca, G. Duke  
; APPLICANT: Martin, Unja  
; APPLICANT: Anderson, Dirk M.  
; TITLE OF INVENTION: CALCIUM/CALMODULIN-DEPENDENT KINASE  
; FILE REFERENCE: 2923-B  
; CURRENT APPLICATION NUMBER: US/10/355,975A  
; CURRENT FILING DATE: 2003-01-30  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 38  
; LENGTH: 630  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; US-10-355-975A-38

Query Match		100.0%;	Score 1342;	DB 4;	Length 630;
Best Local Similarity		100.0%;	Pred. No. 7.7e-130;		
Matches 253;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	YEFLETGKGTGKVKKARESSGR	1VAIKRDKIKDEQDLHRIREIEMSSLNPHI	60	
Db	57	YEFLETGKGTGKVKKARESSGR	1VAIKRDKIKDEQDLHRIREIEMSSLNPHI	116	
Qy	61	IAIHEVFENSKKIVVMEYASRG	LDYDI SERPRLSRDARHFRQIVSALHYCHQNGIV	120	
Db	117	IAIHEVFENSKKIVVMEYASRG	LDYDI SERPRLSRDARHFRQIVSALHYCHQNGIV	176	
Qy	121	HRDLKLENIILLDANGNIKIAD	FGLSNLYHKGKFLQTCGSPLYASPIVNGKPYVGPEVD	180	
Db	177	HRDLKLENIILLDANGNIKIAD	FGLSNLYHKGKFLQTCGSPLYASPIVNGKPYVGPEVD	236	
Qy	181	SWISGLVLLYTLVHGTWPF	DGDHKTIVKOISNGAYREPPKPSDACGLIRLLMVMNPTTRA	240	
Db	237	SWISGLVLLYTLVHGTWPF	DGDHKTIVKOISNGAYREPPKPSDACGLIRLLMVMNPTTRA	296	
Qy	241	TLEDVASHWVWV	253		
Db	297	TLEDVASHWVWV	309		

RESULT 2  
US-09-579-664B-11  
; Sequence 11, Application US/09579664B

```
; Patent No. 6514719
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/09/579,664B
; CURRENT FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-579-664B-11

Query Match 100.0%; Score 1342; DB 4; Length 631;
Best Local Similarity 100.0%; Pred. No. 7.7e-130;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YFLETGKGTGKTYGKVKKARESSGRLVAIKSRKDKIKDEQDLLHIREIEMSSLNHPHI 60
DB 57 YFLETGKGTGKTYGKVKKARESSGRLVAIKSRKDKIKDEQDLLHIREIEMSSLNHPHI 116

QY 61 IAIHEVPENSSKIVIVMEYASRGDLVDYISERPLSRDARHFRQIVSALHYCHQNGIV 120
DB 117 IAIHEVPENSSKIVIVMEYASRGDLVDYISERPLSRDARHFRQIVSALHYCHQNGIV 176

QY 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 180
DB 177 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 236

QY 181 SWSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYRPPKPSDACGLIRLLMWNPTERRA 240
DB 237 SWSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYRPPKPSDACGLIRLLMWNPTERRA 296

QY 241 TLEDVASHWVWVNW 253
DB 297 TLEDVASHWVWVNW 309

RESULT 3
US-10-355-975A-11
; Sequence 11, Application US/10355975A
; Patent No. 6759223
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: CALCIUM/CALMODULIN-DEPENDENT KINASE
; FILE REFERENCE: 2923-B
; CURRENT APPLICATION NUMBER: US/10/355,975A
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-355-975A-11

Query Match 100.0%; Score 1342; DB 4; Length 631;
Best Local Similarity 100.0%; Pred. No. 7.7e-130;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YFLETGKGTGKTYGKVKKARESSGRLVAIKSRKDKIKDEQDLLHIREIEMSSLNHPHI 60
DB 57 YFLETGKGTGKTYGKVKKARESSGRLVAIKSRKDKIKDEQDLLHIREIEMSSLNHPHI 116

QY 61 IAIHEVPENSSKIVIVMEYASRGDLVDYISERPLSRDARHFRQIVSALHYCHQNGIV 120
DB 117 IAIHEVPENSSKIVIVMEYASRGDLVDYISERPLSRDARHFRQIVSALHYCHQNGIV 176

QY 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 180
DB 177 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 236

QY 181 SWSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYRPPKPSDACGLIRLLMWNPTERRA 240
DB 237 SWSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYRPPKPSDACGLIRLLMWNPTERRA 296

QY 241 TLEDVASHWVWVNW 253
DB 297 TLEDVASHWVWVNW 309

RESULT 4
US-08-677-298-2
; Sequence 2, Application US/08677298
; Patent No. 5863729
; GENERAL INFORMATION:
; APPLICANT: Pivnica-Worms, Helen
; TITLE OF INVENTION: DNA SEQUENCES ENCODING HUMAN TCAP-1
; TITLE OF INVENTION: KINASE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/677,298
; FILING DATE: 09-JUL-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Caruthers, Jennie M.
; REGISTRATION NUMBER: 34,464
; REFERENCE/DOCKET NUMBER: 9-96
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 729 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-677-298-2

Query Match 48.4%; Score 649.5; DB 2; Length 729;
Best Local Similarity 51.6%; Pred. No. 3.2e-58;
Matches 131; Conservative 42; Mismatches 78; Indels 3; Gaps 3;

QY 1 YFLETGKGTGKTYGKVKKARESSGRLVAIKSRKDKIKDEQDLLHIREIEMSSLNHPHI 59
DB 56 YRLTKITGKGNFAKVLARHILTGREVAIKIDKTLQ-NPTSLQKLFREVRIMKILNHPN 114

QY 60 IAIHEVPENSSKIVIVMEYASRGDLVDYISERPLSRDARHFRQIVSALHYCHQNGI 119
DB 115 IVKLFEVETETKTYLIMEYASGGEVFDYVAHGRMKREARSKFRQIVSAVQCHOKRI 174

QY 120 VHRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEV 179
DB 175 VHRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEV 234
```



```
QY 239 RATLEDVASHWVN 252
Db 292 RGTLEQIMKDRWN 305

RESULT 8
US-09-523-849-36
; Sequence 36, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scacchi, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523,849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank Accession No. 6458561 gi749794
US-09-523-849-36

Query Match 48.2%; Score 647.5; DB 4; Length 745;
Best Local Similarity 51.2%; Pred. No. 5.3e-58;
Matches 130; Conservative 44; Mismatches 77; Indels 3; Gaps 3;

QY 1 YFELETGKGTGKGVKKARE-SSGRVVAIKSIRKDKIKDEQDLHLHIREIEIMSSLNHPH 59
Db 20 YRLKTTIGKGNFAKVLARHILTGREVAIKIDTQL-NSSSLQKLFREVRIMKVLNHPN 78
QY 60 IIAIHEVFENS SKIVIMVEYASRGDLYDIYISERPRLSERDARHFFRQIVSALHYCHQNGI 119
Db 79 IVKLFVEIETKTLYLIMEYASGGVEFDYLVAGHGRMKEARAKFRQIVSAVOYCHQKFI 138
QY 120 VHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYGPEV 179
Db 139 VHRDLKAENLLLDADNMNIKIADFGSNEFTFGNKLDTFCGSPPYAAPFLFQKKYDGPV 198
QY 180 DWSLGLVLLYLVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDAC-GLIRWLLMVNPT 238
Db 199 DVMSLGVLLYTLVSGSLPFDGQNLKELRERVLRGKIRIPFYMSTDCENLLKFLILNPSK 258
QY 239 RATLEDVASHWVN 252
Db 259 RGTLEQIMKDRWN 272

RESULT 9
US-09-538-092-1022
; Sequence 1022, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01

; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1022
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P27448
US-09-538-092-1022

Query Match 48.2%; Score 646.5; DB 4; Length 724;
Best Local Similarity 50.8%; Pred. No. 6.4e-58;
Matches 129; Conservative 45; Mismatches 77; Indels 3; Gaps 3;

QY 1 YFELETGKGTGKGVKKARE-SSGRVVAIKSIRKDKIKDEQDLHLHIREIEIMSSLNHPH 59
Db 53 YRLKTTIGKGNFAKVLARHILTGREVAIKIDTQL-NSSSLQKLFREVRIMKVLNHPN 111
QY 60 IIAIHEVFENS SKIVIMVEYASRGDLYDIYISERPRLSERDARHFFRQIVSALHYCHQNGI 119
Db 112 IVKLFVEIETKTLYLIMEYASGGVEFDYLVAGHGRMKEARAKFRQIVSAVOYCHQKFI 171
QY 120 VHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYGPEV 179
Db 172 VHRDLKAENLLLDADNMNIKIADFGSNEFTFGNKLDTFCGSPPYAAPFLFQKKYDGPV 231
```





Db 112 IVKLEFVETETKTLVMEYASGGEVFDYLVAGHGMKEKARAKFRQIVSAVQCHHKFI 171  
Qy 120 VHRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPIVNGKPYVGPEV 179  
Db 172 VHRDLKAENLLDDAMNIKIADFGSNEFTGCKLDITFCGSPPYAAPFQGGKYDGPV 231  
Qy 180 DWSLSGLVLLYTLVHGTPFPDQDHKTLVKQISNGAYREPPKPSDAC-GLIRWLLMVNPT 238  
Db 232 DVWSLGVILYTLVSGSLPFDQNLKELRERVLRGKYRIPFYVMSTDCENLLKKLILNPSK 291  
Qy 239 RATLEDVASHWVN 252  
Db 292 RGSLEQIMKORWMN 305

RESULT 14  
US-09-523-849-32  
; Sequence 32, Application US/09523849  
; Patent No. 6458561  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Molteni, Angela  
; APPLICANT: Magnaghi, Paola  
; APPLICANT: Bosotti, Roberta  
; APPLICANT: Scaccheri, Emanuela  
; APPLICANT: Isacchi, Antonella  
; APPLICANT: Hodgson, Dave  
; TITLE OF INVENTION: HUMAN NTM1 KINASE  
; FILE REFERENCE: PC-0009 US  
; CURRENT APPLICATION NUMBER: US/09/523,849  
; CURRENT FILING DATE: 2000-03-13  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PERL Program  
; SEQ ID NO 32  
; LENGTH: 793  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: GenBank Accession No. 6458561 g2052189  
US-09-523-849-32

Query Match 47.9%; Score 642.5; DB 4; Length 793;  
Best Local Similarity 51.2%; Pred. No. 1.9e-57;  
Matches 130; Conservative 43; Mismatches 78; Indels 3; Gaps 3;  
Qy 1 YEFLETGKGTGKGVKKARE-SSGRLVAIKSIRKDKIKDEQDLHHRREIEIMSSLNHPH 59  
Db 60 YRLQKTIGKGNFAVKLARHVLTCREAVAKIIDTKQL-NPTSLQKLFREVRIMKILNPN 118  
Qy 60 IIAIHEVFENSSKIVIMEYASRGDLYISERPRLSERDARHFRQIVSALHYCHONGI 119  
Db 119 IVKLEFVETETKTLVMEYASGGEVFDYLVAGHGMKEKARAKFRQIVSAVQCHQKCI 178  
Qy 120 VHRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPIVNGKPYVGPEV 179  
Db 179 VHRDLKAENLLDDAMNIKIADFGSNEFTGCKLDITFCGSPPYAAPFQGGKYDGPV 238  
Qy 180 DWSLSGLVLLYTLVHGTPFPDQDHKTLVKQISNGAYREPPKPSDAC-GLIRWLLMVNPT 238  
Db 239 DVWSLGVILYTLVSGSLPFDQNLKELRERVLRGKYRIPFYVMSTDCENLLKKLILNPIK 298  
Qy 239 RATLEDVASHWVN 252  
Db 299 RGSLEQIMKORWMN 312

RESULT 15  
US-09-633-328B-4  
; Sequence 4, Application US/09633328B  
; Patent No. 6777587  
; GENERAL INFORMATION:

; APPLICANT: BISARO, DAVID A.  
; TITLE OF INVENTION: METHOD OF ENHANCING PLANT RESISTANCE TO PATHOGENS  
; FILE REFERENCE: 22727-04041  
; CURRENT APPLICATION NUMBER: US/09/633,328B  
; CURRENT FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: 60/147,613  
; PRIOR FILING DATE: 1999-08-06  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patentin Ver. 3.2  
; SEQ ID NO 4  
; LENGTH: 511  
; TYPE: PRT  
; ORGANISM: Nicotiana benthamiana  
US-09-633-328B-4

Query Match 47.0%; Score 631; DB 4; Length 511;  
Best Local Similarity 49.6%; Pred. No. 1.6e-56;  
Matches 125; Conservative 48; Mismatches 77; Indels 2; Gaps 2;  
Qy 1 YEFLETGKGTGKGVKKARES-SGRLVAIKSIRKDKIKDEQDLHHRREIEIMSSLNHPH 59  
Db 19 YKLGKTLIGSGFGKVKIAEHTLTGHKVAVKILNRRKIKNMEMEKEKVRREIKILFLFWHPH 78  
Qy 60 IIAIHEVFENSSKIVIMEYASRGDLYISERPRLSERDARHFRQIVSALHYCHONGI 119  
Db 79 IIRLYEVVETPSDIYVMEYVVKSGELFDYIVKGRLOEDEARKEFQIISGVEYCHRNWV 138  
Qy 120 VHRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPIVNGKPYVGPEV 179  
Db 139 VHRDLKLENLLDSKMNKVIADFGLSNIMRDGHFLKTCGSPNYAAPEVTSGLYAGPEV 198  
Qy 180 DWSLSGLVLLYTLVHGTPFPDQDHKTLVKQISNGAYREPPKPS-DACGLIRWLLMVNPT 238  
Db 199 DVWSGVLVYLLCGTLFPDDENIPNLFKKIKGMISLPSHLSAGARDLIPRMLIVDPMK 258  
Qy 239 RATLEDVASHWV 250  
Db 259 RMTIPEIRMHWP 270

Search completed: January 22, 2005, 06:44:55  
Job time : 17.3133 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 06:22:58 ; Search time 63.8224 Seconds  
(without alignments)  
1422.048 Million cell updates/sec

Title: US-09-980-464-11\_COPY\_57\_309

Perfect score: 1342

Sequence: 1 YEFLETGLKGYGVKKARE.....VNPTTRATLEDVASHWVNVW 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04.\*  
1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1342	100.0	631	4 AAB50056	Aab50056 Murine Ly
2	1313	97.8	630	5 AAE19885	Aae19885 Rat SNF1/
3	1293	96.3	628	4 AAM93360	Aam93360 Human pol
4	1293	96.3	628	4 ABU53319	Abu53319 Human cel
5	1293	96.3	628	5 ABP69116	Abp69116 Human pol
6	1293	96.3	628	5 AAU79652	Aau79652 Human pro
7	1293	96.3	628	7 ADF76965	Adf76965 Novel hum
8	1293	96.3	628	8 ADL30886	Adl30886 Human pro
9	1293	96.3	628	8 ADL25362	Adl25362 Human SNA
10	1293	96.3	628	8 ADL14161	Adl14161 Novel hum
11	1293	96.3	628	8 ADL20172	Adl20172 Human PRO
12	1293	96.3	672	8 ADJ96620	Adj96620 Human cal
13	1286	95.8	594	5 AAE16266	Aae16266 Human kin
14	1279.5	95.3	611	6 ABP96085	Abp96085 Human pro
15	1268.5	94.5	629	4 AAB65632	Aab65632 Novel pro
16	1268.5	94.5	629	8 ADL29239	Adl29239 Human MAR
17	1128.5	84.1	660	7 ADN95766	Adn95766 Human BEC
18	1128.5	84.1	661	4 AAE07847	Aae07847 Human pro
19	1128.5	84.1	661	4 AAE07846	Aae07846 Human pro
20	1128.5	84.1	661	7 ADE38421	Ade38421 Human pro
21	1128.5	84.1	661	8 ADJ75331	Adj75331 Marker ge
22	1128.5	84.1	661	8 ADL25353	Adl25353 Human ARK
23	1128.5	84.1	661	8 ADQ19734	Adq19734 Human sof
24	1085	80.8	534	4 AAB71959	Aab71959 Human TGF
25	822.5	61.3	434	4 ABB67451	Abb67451 Drosophil

26	814	60.7	530	4 AAE00668	Aae00668 Human pro
27	665.5	49.6	508	5 ABG79179	Abg79179 Human ser
28	665.5	49.6	520	4 AAB65629	Aab65629 Novel pro
29	665.5	49.6	520	8 ADI29235	Adi29235 Mouse MAR
30	665.5	49.6	688	5 ABB04434	Abb04434 Human neu
31	665.5	49.6	689	7 ADF74129	Adf74129 Human nov
32	665.5	49.6	752	4 AAM93956	Aam93956 Human pol
33	665.5	49.6	752	4 AAE11782	Aae11782 Human kin
34	665.5	49.6	752	5 ABB04433	Abb04433 Human neu
35	665.5	49.6	752	6 ABG73794	Abg73794 Human MAR
36	665.5	49.6	752	8 ADL32125	Adl32125 Human pro
37	665.5	49.6	769	5 ABP62966	Abp62966 Human pol
38	661.5	49.3	752	8 ADJ96622	Adj96622 Human cal
39	660.5	49.2	660	5 ABB04432	Abb04432 Murine ne
40	660.5	49.2	688	6 AAE33554	Aae33554 Human mic
41	660.5	49.2	688	6 AAE33555	Aae33555 Human mic
42	660.5	49.2	688	7 ADG91726	Adg91726 Human mic
43	660.5	49.2	688	7 ADG91727	Adg91727 Human mic
44	660.5	49.2	724	5 ABB04431	Abb04431 Murine ne
45	653	48.7	761	6 ABP96068	Abp96068 Human pro

ALIGNMENTS

RESULT 1  
AAB50056  
ID AAB50056 standard; protein; 631 AA.  
XX AAB50056;

AC AC  
XX XX  
DT 19-MAR-2001 (first entry)  
XX Murine Lymph node Stromal cell kinase 1.  
DE  
XX Murine; Lymph node Stromal cell kinase; MLCK-1; autoimmune disorder;  
KW wound healing; periodontal disease; inflammatory disease; tumour;  
KW infection; allergy.  
XX  
OS Mus musculus.  
XX  
XX WO200073468-Al.  
XX  
PD 07-DEC-2000.  
XX  
XX 26-MAY-2000; 2000WO-US014696.  
XX  
XX 28-MAY-1999; 99US-0136781P.  
XX  
XX (IMMV ) IMMUNEX CORP.  
XX  
XX Bird TA, Virca GD, Martin U, Anderson DM;  
XX  
XX WPI; 2001-061546/07.  
XX  
XX N-PSDB; AAC90433.  
XX  
XX Novel murine and human kinase nucleic acids useful for treating  
XX inflammations, infections, tumors, allergies, autoimmune diseases, and  
XX for stimulating or suppressing immune responses.  
XX  
XX Claim 10; Page 94-96; 106pp; English.  
XX  
XX The present sequence is Murine Lymph node Stromal cell kinase 1 (MLSK-1).  
XX This protein is useful for treating a variety of disorders listed in the  
XX disclosure of the specification, including autoimmune disorders, allergic  
XX reactions, myeloid or lymphoid cell deficiencies, wound healing and  
XX tissue repair and replacement, burns, incisions and ulcers, periodontal  
XX disease, inflammatory diseases, tumours and bacterial, viral or fungal  
XX infection  
XX  
SQ Sequence 631 AA;

Query Match 100.0%; Score 1342; DB 4; Length 631;

Best Local Similarity 100.0%; Pred. No. 2.6e-138; Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YEFLTLGKTYGKVKKARESSGRLVAIKSRKDKIKDEQDLHRRREIEMSLNHPHI 60  
 Db 57 YEFLTLGKTYGKVKKARESSGRLVAIKSRKDKIKDEQDLHRRREIEMSLNHPHI 116  
 Qy 61 IAIHEVFENSSKIVIVMEYASRGDLVDYISERPLSERDARHFRQIVSALHYCHQNGIV 120  
 Db 117 IAIHEVFENSSKIVIVMEYASRGDLVDYISERPLSERDARHFRQIVSALHYCHQNGIV 176  
 Qy 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGFQTFGCSPLASPEIVNGKPYVGVPEVD 180  
 Db 177 HRDLKLENILLDANGNIKIADFGLSNLYHKGFQTFGCSPLASPEIVNGKPYVGVPEVD 236  
 Qy 181 SWSLGVLLYLIVHGTMPFGDQDKHTLVKQISNGAYREPPKPSDACGLIRLLMVPTRRA 240  
 Db 237 SWSLGVLLYLIVHGTMPFGDQDKHTLVKQISNGAYREPPKPSDACGLIRLLMVPTRRA 296  
 Qy 241 TLEDVASHWVNW 253  
 Db 297 TLEDVASHWVNW 309

RESULT 2  
 AA019885  
 ID AA019885 standard; protein; 630 AA.  
 AC AA019885;  
 XX  
 XX  
 DT 18-JUN-2002 (first entry)  
 XX  
 DE Rat SNF1/AMPK-Related Kinase (SNARK) protein.  
 XX  
 KW Rat; SNF1/AMPK-Related Kinase; SNARK; enzyme; stress response; diabetes;  
 KW glucose deprivation; lipid metabolism; therapy; lipoprotein disorder;  
 KW hyperglycaemic; drug screening; hypoglycaemia.  
 XX  
 OS Rattus sp.  
 XX

Key Location/Qualifiers  
 Binding-site 63..89  
 FT /note= "Protein kinase ATP-binding region signature"  
 FT 137..140  
 FT /note= "Serine/threonine kinase catalytic domain"  
 FT /note= "Serine/threonine kinase catalytic domain"  
 FT 175..187  
 FT /note= "Serine/threonine protein kinase active-site"  
 FT 297..300  
 FT /note= "Serine/threonine kinase catalytic domain"  
 FT 335..338  
 FT /note= "Serine/threonine kinase catalytic domain"  
 FT 381..384  
 FT /note= "Serine/threonine kinase catalytic domain"  
 FT 422..425  
 FT /note= "Serine/threonine kinase catalytic domain"  
 FT 468..471  
 FT /note= "Serine/threonine kinase catalytic domain"  
 FT 517..520  
 FT /note= "Serine/threonine kinase catalytic domain"  
 FT 601..604  
 FT /note= "Serine/threonine kinase catalytic domain"  
 FT 608..611  
 FT /note= "Serine/threonine kinase catalytic domain"  
 XX  
 PN W0200212456-A2.  
 XX  
 PD 14-FEB-2002.  
 XX  
 XX 02-AUG-2001; 2001WO-CA001109.  
 XX  
 PR 03-AUG-2000; 2000US-0222650P.  
 PR 12-MAR-2001; 2001US-0274613P.  
 PR 28-MAR-2001; 2001CA-02340783.

XX  
 PA (ONEO-) 1149336 ONTARIO INC.  
 XX  
 PT Drucker DJ, Rosen CF, Lefebvre DL;  
 XX  
 DR WPI; 2002-241747/29.  
 DR N-PSDB; AAD31710.  
 XX  
 XX AMPK (AMP-activated protein kinase)-related kinase, designated SNARK  
 FT polypeptides and polynucleotides, useful for treating or preventing  
 FT diabetes, or other disorders of lipoprotein production leading to  
 FT increased levels of cholesterol.  
 XX  
 XX Claim 1; Fig 2; 94pp; English.  
 PS  
 XX  
 CC The invention relates to an AMPK (AMP-activated protein kinase)-related  
 CC kinase, designated SNARK polypeptides and polynucleotides. SNARK  
 CC (SNIP/AMP-activated protein kinase) is involved in stress response to  
 CC glucose deprivation. The polynucleotides are useful for expressing SNARK  
 CC protein in isolated form or as a protein conjugate. Activation of SNARK  
 CC stimulates liver CPT-1 thus enhances lipid metabolism in liver cells and  
 CC in other cell types such as heart and skeletal muscles, as well as  
 CC increases GLUT-4 and glycogen in muscle. Activation of SNARK is predicted  
 CC to have insulin-like effects that would enhance the disposal of glucose  
 CC into muscle and reduce plasma glucose for the treatment of diabetes and  
 CC some type of disorders of lipoprotein production leading to increased  
 CC levels of cholesterol or triglycerides. SNARK or its variants may be  
 CC administered to a subject to treat or prevent a disease associated with  
 CC decreased expression of SNARK, such as diabetes. SNARK antibodies are  
 CC used to modulate SNARK activity either in vivo for therapeutic purposes,  
 CC or in vitro for drug screening and related investigational purposes.  
 CC SNARK antagonists may be administered to increase fuel production,  
 CC decrease glucose uptake and increase levels of blood glucose in a patient  
 CC suffering from hypoglycaemia. The present sequence is rat SNARK protein  
 XX  
 SQ Sequence 630 AA;

Query Match 97.8%; Score 1313; DB 5; Length 630;  
 Best Local Similarity 98.0%; Pred. No. 4.1e-135;  
 Matches 248; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YEFLTLGKTYGKVKKARESSGRLVAIKSRKDKIKDEQDLHRRREIEMSLNHPHI 60  
 Db 57 YEFLTLGKTYGKVKKARESSGRLVAIKSRKDKIKDEQDLHRRREIEMSLNHPHI 116  
 Qy 61 IAIHEVFENSSKIVIVMEYASRGDLVDYISERPLSERDARHFRQIVSALHYCHQNGIV 120  
 Db 117 IAIHEVFENSSKIVIVMEYASRGDLVDYISERPLSERDARHFRQIVSALHYCHQNGIV 176  
 Qy 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGFQTFGCSPLASPEIVNGKPYVGVPEVD 180  
 Db 177 HRDLKLENILLDANGNIKIADFGLSNLYHKGFQTFGCSPLASPEIVNGKPYVGVPEVD 236  
 Qy 181 SWSLGVLLYLIVHGTMPFGDQDKHTLVKQISNGAYREPPKPSDACGLIRLLMVPTRRA 240  
 Db 237 SWSLGVLLYLIVHGTMPFGDQDKHTLVKQISNGAYREPPKPSDACGLIRLLMVPTRRA 296  
 Qy 241 TLEDVASHWVNW 253  
 Db 297 TLEDVASHWVNW 309

RESULT 3  
 AA093360  
 ID AA093360 standard; protein; 628 AA.  
 XX  
 XX AC AA093360;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human polypeptide, SEQ ID NO: 2919.  
 XX  
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.

XX OS Homo sapiens.  
XX PN EP1130094-A2.  
XX PD 05-SEP-2001.  
XX PF 07-JUL-2000; 2000EP-00114089.  
XX PR 08-JUL-1999; 99JP-00194486.  
XX PR 11-JAN-2000; 2000JP-00118774.  
XX PR 02-MAY-2000; 2000JP-00183785.  
XX PA (HELI-) HELIX RES INST.  
XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX DR N-PSDB; AAK94280.  
XX PS WPI; 2001-524255/58.  
XX PT 830 Primers useful for synthesizing full length cDNA clones and their use  
PT in genetic manipulation.  
XX PS Claim 8; SEQ ID NO 2919; 1380pp + Sequence Listing; English.  
XX CC The invention relates to primers for synthesizing full length cDNA  
CC clones. 830 cDNA molecules encoding a human protein have been isolated  
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have  
CC been determined. Primers for synthesizing the full length cDNA are useful  
CC for clarifying the function of the protein encoded by the cDNA. The full  
CC length clones were obtained by construction of full length enriched cDNA  
CC libraries that were synthesised by the oligo-capping method. The primers  
CC enable the production of the full length cDNA easily without any special  
CC methods. The present sequence is a polypeptide encoded by a full length  
CC human cDNA of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in CD-ROM  
CC format directly from EPO  
XX SQ Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 4; Length 628;  
Best Local Similarity 95.7%; Pred. No. 6.5e-133;  
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 YEFLETGKGTGKVKKARESGRLVAIKSRKDKIKDEQDLHRRREIETMSSLNHPHI 60  
Db 53 YEFLETGKGTGKVKKARESGRLVAIKSRKDKIKDEQDLHRRREIETMSSLNHPHI 112  
Qy 61 IAIHEVFENSSKIVIVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 120  
Db 113 IAIHEVFENSSKIVIVMEYASRGDLVDYISERQQLSEREARRHFRQIVSAVHYCHONRVV 172  
Qy 121 HRDLKLENILLDANGNIKIADFGLSNLVYHKGKFLQTFCGSPLYASPEIVNGKPYVGPEVD 180  
Db 173 HRDLKLENILLDANGNIKIADFGLSNLVYHKGKFLQTFCGSPLYASPEIVNGKPYTGPEVD 232  
Qy 181 SWSLGVLLYLIVHGTMPEFDGQDHKTLVKQISNGAYREPPEKPSDACGLIRWLLMVPNPTTRA 240  
Db 233 SWSLGVLLYLIVHGTMPEFDGHDHKLIVKQISNGAYREPPEKPSDACGLIRWLLMVPNPTTRA 292  
Qy 241 TLEDVASHHWVNW 253  
Db 293 TLEDVASHHWVNW 305

RESULT 4  
ABU53319  
ID ABU53319 standard; protein; 628 AA.  
XX AC ABU53319;  
XX DT 14-APR-2003 (first entry)

XX DE Human cell cycle-associated protein from DKFZphtes3\_7f3.  
XX KW Human; gene therapy; vaccine; disease treatment; detection.  
XX OS Homo sapiens.  
XX PN WO200112659-A2.  
XX PD 22-FEB-2001.  
XX PF 18-AUG-2000; 2000WO-IB001496.  
XX PR 18-AUG-1999; 99US-0149499P.  
XX PR 28-SEP-1999; 99US-0156503P.  
XX PA (GEHU-) GERMAN HUMAN GENOME PROJECT.  
XX PI Wiemann S;  
XX PI WPI; 2001-327840/34.  
XX DR N-PSDB; ABX71420.  
XX PS Nucleic acids having the sequences of clones isolated from libraries of  
PT different human tissues, useful in recombinant DNA methodologies.  
XX PS Claim 21; Page 943; 1095pp; English.  
XX CC This invention describes novel polynucleotides and polypeptides isolated  
CC from human cDNA libraries which can be used for gene therapy or in  
CC vaccines. The polynucleotides of the invention and antibodies encoded by  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate polypeptide expression. The products of the  
CC invention may also be used to identify modulators of expression and  
CC activity and to down regulate expression and activity. The antibodies of  
CC the invention may also be used as diagnostic agents for detecting the  
CC presence of polypeptides in samples. This sequence represents a  
CC polypeptide described in the disclosure of the invention  
XX SQ Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 4; Length 628;  
Best Local Similarity 95.7%; Pred. No. 6.5e-133;  
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 YEFLETGKGTGKVKKARESGRLVAIKSRKDKIKDEQDLHRRREIETMSSLNHPHI 60  
Db 53 YEFLETGKGTGKVKKARESGRLVAIKSRKDKIKDEQDLHRRREIETMSSLNHPHI 112  
Qy 61 IAIHEVFENSSKIVIVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 120  
Db 113 IAIHEVFENSSKIVIVMEYASRGDLVDYISERQQLSEREARRHFRQIVSAVHYCHONRVV 172  
Qy 121 HRDLKLENILLDANGNIKIADFGLSNLVYHKGKFLQTFCGSPLYASPEIVNGKPYVGPEVD 180  
Db 173 HRDLKLENILLDANGNIKIADFGLSNLVYHKGKFLQTFCGSPLYASPEIVNGKPYTGPEVD 232  
Qy 181 SWSLGVLLYLIVHGTMPEFDGQDHKTLVKQISNGAYREPPEKPSDACGLIRWLLMVPNPTTRA 240  
Db 233 SWSLGVLLYLIVHGTMPEFDGHDHKLIVKQISNGAYREPPEKPSDACGLIRWLLMVPNPTTRA 292  
Qy 241 TLEDVASHHWVNW 253  
Db 293 TLEDVASHHWVNW 305

RESULT 5  
ABP69116  
ID ABP69116 standard; protein; 628 AA.  
XX AC ABP69116;  
XX DT 20-JAN-2003 (first entry)

XX DE Human polypeptide SEQ ID NO 1163.

XX KW Human; genome mapping; gene therapy; food supplement; virus; fungus;

XX KW cell-proliferative disorder; neurodegenerative disease; bacterial;

XX KW Parkinson's disease; Alzheimer's disease; autoimmune disease;

XX KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;

XX KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;

XX KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;

XX KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;

XX KW antiarthritic.

XX OS Homo sapiens.

XX PN WO200270539-A2.

XX PD 12-SEP-2002.

XX PF 05-MAR-2002; 2002WO-US005095.

XX PR 05-MAR-2001; 2001US-00799451.

XX FA (HYSE-) HVSEQ INC.

XX PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;

XX PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;

XX PI Wehrman T, Wang J, Wang D, Drmanac RT;

XX DR WPI; 2002-759812/82.

XX DR N-PSDB; ABZ11333.

XX PT New polynucleotides comprising sequences assembled from expressed

XX PT sequence tags (ESTs), useful for treating cell-proliferative,

XX PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet

XX PT or coagulation disorders.

XX PS Claim 9; SEQ ID NO 1163; 1012pp + Sequence Listing; English.

XX CC The invention relates to an isolated polynucleotide (I) comprising a

XX CC nucleotide sequence selected from any of 948 sequences (ABZ11119-

XX CC ABZ12066) or their mature protein coding portion, active domain coding

XX CC protein or complementary sequences. The polynucleotides are useful for

XX CC identifying expressed genes or for physical mapping of human genome. The

XX CC encoded polypeptides (ABP68902-ABP6949) are useful as molecular weight

XX CC markers, as a food supplement, for generating antibodies, in medical

XX CC imaging, screening and diagnostic assays and for treating cell-

XX CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's

XX CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,

XX CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,

XX CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver

XX CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),

XX CC arthritis, etc. Note: The sequence data for this patent did not form part

XX CC of the printed specification, but was obtained in electronic format

XX CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 5; Length 628;

Best Local Similarity 95.7%; Pred. No. 6.5e-133;

Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YEFLETGKGYGVKKARESSGRLVAIKSRKDKIKDEQDLHIREIETMSLNPHI 60

DB 53 YEFLETGKGYGVKKARESSGRLVAIKSRKDKIKDEQDLHIREIETMSLNPHI 112

QY 61 IAIHEVFENSKTIVVMEYASRGDLVYISRPRLSERDARHFRQIVSALHYCHQNGIV 120

DB 113 IAIHEVFENSKTIVVMEYASRGDLVYISERQOLSERARHFRQIVSAVHYCHQNRV 172

QY 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPIVNGKPYVGPEVD 180

DB 173 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPIVNGKPYVGPEVD 232

QY 181 SWSLGVLLYILVHCTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRLLMVNPTTRA 240

DB 233 SWSLGVLLYILVHCTMPFDGHDHKLKLVKQISNGAYREPPKPSDACGLIRLLMVNPTTRA 292

QY 241 TLEDVASHWVNVN 253

DB 293 TLEDVASHWVNVN 305

RESULT 6

AAU79652

ID AAU79652 standard; protein; 628 AA.

XX AC AAU79652;

XX DT 02-JUL-2002 (first entry)

XX DE Human protein kinase 3700.

XX KW Human; protein kinase 3700; PK; protein phosphorylation; tumorigenesis;

XX KW cell signalling; mitogenesis; gene transcription; angiogenesis; sarcoma;

XX KW tissue repair; tissue regeneration; atherosclerosis; blood-brain barrier;

XX KW cell proliferation disorder; cell differentiation disorder; carcinoma;

XX KW haematopoietic neoplastic disorder; metastatic disorder; leukaemia;

XX KW cytostatic; antiatherosclerotic; enzyme.

XX OS Homo sapiens.

XX PN WO200224921-A2.

XX PD 28-MAR-2002.

XX PF 25-SEP-2001; 2001WO-US030115.

XX PR 25-SEP-2000; 2000US-0234922P.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Curtis RAJ, Galvin KM;

XX WPI; 2002-352007/38.

XX N-PSDB; ABK14000.

XX PT Use of modulators of activity of 3700 protein for making medicament for

XX PT e.g., modulating protein phosphorylation or cell signaling, or for

XX PT treating or preventing cellular proliferative and/or differentiative

XX PT disorders.

XX PS Claim 19; Fig 1; 115pp; English.

XX CC The present invention relates to the isolation of a novel human protein

XX CC kinase designated 3700, and the polynucleotide sequence encoding it. The

XX CC invention also describes the use of a modulator of the activity of

XX CC protein kinase (PK) 3700 for making a medicament or pharmaceutical

XX CC composition for modulating the ability of a cell to phosphorylate an

XX CC amino acid residue of a substrate protein. Modulators of protein kinase

XX CC 3700 activity are useful for modulating protein phosphorylation, cell

XX CC signalling, tumorigenesis, mitogenesis, transcription of a gene,

XX CC angiogenesis, tissue repair, tissue regeneration, establishment or

XX CC progression of atherosclerosis, and signalling across the blood-brain

XX CC barrier. The polynucleotide and polypeptide molecules for protein kinase

XX CC 3700 may be used as diagnostic targets and therapeutic agents for

XX CC prognosticating, diagnosing, preventing, inhibiting, alleviating, or

XX CC curing PK-related disorders and cellular proliferative and/or

XX CC differentiative disorders (e.g. haematopoietic neoplastic disorders,

XX CC carcinoma, sarcoma, metastatic disorders or leukaemia). The 3700

XX CC polynucleotide sequence can be used to express protein kinase 3700, to

XX CC detect a genetic alteration in a 3700 gene, in chromosome mapping, for

XX CC tissue typing, in forensic biology, and as surrogate markers. The present

XX CC sequence represents human protein kinase 3700

XX SQ Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 5; Length 628;  
Best Local Similarity 95.7%; Pred. No. 6.5e-133;  
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YEFLETGKGYGVKKKARESSGRLVAIKSRDKIKDEODLLHRRREIETMSSLNHPHI 60  
DB 53 YEFLETGKGYGVKKKARESSGRLVAIKSRDKIKDEODLLHRRREIETMSSLNHPHI 112  
QY 61 IAIHEVFENSKIVIVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHQNGIV 120  
DB 113 IAIHEVFENSKIVIVMEYASRGDLVDYISERQQLSEREARHFRQIVSAVHYCHQNRVV 172  
QY 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 180  
DB 173 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 232  
QY 181 SWSLGVLLYILVHGTMPDFGDHKTIVKQISNGAYRPPKPSDACGLIRWLLMVPTRRA 240  
DB 233 SWSLGVLLYILVHGTMPDFGDHKTIVKQISNGAYRPPKPSDACGLIRWLLMVPTRRA 292  
QY 241 TLEDVASHWVNW 253  
DB 293 TLEDVASHWVNW 305

RESULT 7  
ADF76965  
ID ADF76965 standard; protein; 628 AA.  
XX  
AC ADF76965;  
XX  
XX 26-FEB-2004 (first entry)  
XX  
XX Novel human secreted and transmembrane protein SeqID 640.  
XX human; PRO; membrane bound protein; membrane bound receptor;  
KW cell proliferation; cell migration; cell differentiation;  
KW mitogenic factor; survival factor; cytotoxic factor;  
KW differentiation factor; neuroepithelial; hormone; cell receptor;  
KW receptor-ligand interaction; cytostatic; chondrocyte; tumour.  
XX  
OS Homo sapiens.  
XX  
XX WO2003072035-A2.  
XX  
XX 04-SEP-2003.  
XX  
XX 21-FEB-2003; 2003WO-US005241.  
XX  
XX 22-FEB-2002; 2002US-0359461P.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;  
PI Williams PM, Wood WI, Wu TD;  
XX  
XX WPI; 2003-721702/68.  
XX  
XX N-PSDB; ADF76964.  
XX  
XX New PRO polypeptides, useful for diagnosing and treating an immune  
PT related disorder, e.g. systemic lupus erythematosus, rheumatoid  
PT arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or  
PT diabetes mellitus.  
XX  
XX Claim 10; SEQ ID NO 640; 918pp; English.  
XX  
XX This invention relates to novel nucleic acids encoding human PRO secreted  
CC and transmembrane proteins. Extracellular proteins play important roles  
CC in the formation, differentiation and maintenance of multicellular  
CC organisms. The fate of many individual cells (for example proliferation,  
CC migration or differentiation) is typically governed by information  
CC received from other cells and the immediate environment. The information  
CC is often transmitted by secreted polypeptides (for example mitogenic

factors, survival factors, cytotoxic factors, differentiation factors, neuroepithelial factors (hormones) which are received and interpreted by diverse cell receptors or membrane bound proteins. These membrane bound proteins and receptors may be of use as pharmaceutical and diagnostic agents, such as in the blocking of receptor-ligand interactions. The current invention provides the amino acid sequences of novel human membrane bound receptors and proteins, along with the cDNA sequences encoding them. The novel proteins of the invention may have cytostatic activities through the stimulation of chondrocytes. The nucleic acids of the invention may be useful for the manufacture of a medicament for diagnosing or treating a tumour in a mammal. In addition, they may be useful for measuring or detecting the expression of a tumour associated gene. The present sequence is the amino acid sequence of a human PRO protein of the invention.

Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 7; Length 628;  
Best Local Similarity 95.7%; Pred. No. 6.5e-133;  
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YEFLETGKGYGVKKKARESSGRLVAIKSRDKIKDEODLLHRRREIETMSSLNHPHI 60  
DB 53 YEFLETGKGYGVKKKARESSGRLVAIKSRDKIKDEODLLHRRREIETMSSLNHPHI 112  
QY 61 IAIHEVFENSKIVIVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHQNGIV 120  
DB 113 IAIHEVFENSKIVIVMEYASRGDLVDYISERQQLSEREARHFRQIVSAVHYCHQNRVV 172  
QY 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 180  
DB 173 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 232  
QY 181 SWSLGVLLYILVHGTMPDFGDHKTIVKQISNGAYRPPKPSDACGLIRWLLMVPTRRA 240  
DB 233 SWSLGVLLYILVHGTMPDFGDHKTIVKQISNGAYRPPKPSDACGLIRWLLMVPTRRA 292  
QY 241 TLEDVASHWVNW 253  
DB 293 TLEDVASHWVNW 305

RESULT 8  
ADL30886  
ID ADL30886 standard; protein; 628 AA.  
XX  
AC ADL30886;  
XX  
XX 20-MAY-2004 (first entry)  
XX  
XX Human protein encoded by a full length cDNA clone SeqID 2919.  
XX human; medicine; signal transduction; glycoprotein; transcription;  
KW oligo-capping method.  
XX  
OS Homo sapiens.  
XX  
XX EP1396543-A2.  
XX  
XX 10-MAR-2004.  
XX  
XX 07-JUL-2000; 2003EP-00025638.  
XX  
XX 08-JUL-1999; 99JP-00194486.  
PR 11-JAN-2000; 2000JP-0018774.  
PR 02-MAY-2000; 2000JP-00183865.  
PR 07-JUL-2000; 2000EP-00114089.  
XX  
XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
XX Oka T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX

DR WPI; 2004-204755/20.  
 DR N-PSDB; ADL30885.  
 XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full  
 PT length human cDNAs.  
 XX Example 1; SEQ ID NO 2919; 1340pp; English.  
 PS This invention relates to a novel primers useful for synthesizing full  
 CC length cDNA molecules that encode human proteins. Specifically, it refers  
 CC to secretory or membrane proteins that are potential therapeutic agents/  
 CC target molecules in the field of medicine, and in particular genes  
 CC encoding proteins that are associated with signal transduction,  
 CC glycoproteins and transcription. The present invention describes a method  
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'  
 CC ends using the oligo-capping method. This polypeptide sequence is a full  
 CC length human protein of the invention.  
 XX  
 SQ Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 8; Length 628;  
 Best Local Similarity 95.7%; Pred. No. 6.5e-133;  
 Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 YEFLETIGKGTGKVKKARESSGRLVAIKSIRKDKIKDEQDLHRRREIEMSSLNHPHI 60  
 DB 53 YEFLETIGKGTGKVKKARESSGRLVAIKSIRKDKIKDEQDLHRRREIEMSSLNHPHI 112  
 QY 61 IAIHEVFENSKIVIVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 120  
 DB 113 IAIHEVFENSKIVIVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 172  
 QY 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPVVGPEVD 180  
 DB 173 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPVVGPEVD 232  
 QY 181 SWSGLVLLYLIVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDAGLIRLLMVNPTERRA 240  
 DB 233 SWSGLVLLYLIVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDAGLIRLLMVNPTERRA 292  
 QY 241 TLEDVASHWVWVW 253  
 DB 293 TLEDVASHWVWVW 305

RESULT 9  
 ADL25362  
 ID ADL25362 standard; protein; 628 AA.  
 XX  
 AC ADL25362;  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Human SNARK, SEQ ID 12.  
 XX  
 KW Cytostatic; Neuroprotective; Nootropic; Muscular; Gene therapy;  
 KW Adenosine Monophosphate activated Protein Kinase-related kinase 5;  
 KW AMPK-related kinase 5; ARK5; KIAA0537; stress resistance; tumour;  
 KW nervous disorder; muscle disorder; ataxia teleangiectasia; SNARK.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2004019994-A1.  
 XX  
 XX 11-MAR-2004.  
 XX  
 XX 19-AUG-2003; 2003WO-JP010435.  
 XX  
 XX 27-AUG-2002; 2002JP-00247761.  
 XX  
 XX (NINA-) JAPAN AGENCY NAT INST HEALTH.  
 XX (PHAR-) ORG PHARM SAFETY & RES.  
 XX

PI Esumi H, Suzuki A;  
 XX  
 DR WPI; 2004-248195/23.  
 XX  
 PT Agent for imparting stress resistance to cells, comprises DNA encoding  
 PT for adenosine monophosphate activated protein kinase-related kinase 5 for  
 PT treatment of tumors, muscle and nervous disorders, and ataxia  
 PT teleangiectasia.  
 XX  
 XX Example 1; SEQ ID NO 12; 143pp; Japanese.  
 PS The present invention relates to human Adenosine Monophosphate activated  
 CC Protein Kinase (AMPK)-related kinase 5 (ARK5/ KIAA0537) and its coding  
 CC sequence, which are useful for imparting stress resistance to cells. ARK5  
 CC is useful for treatment and prevention of tumours, nervous disorders,  
 CC muscle disorders and ataxia teleangiectasia. The present sequence was  
 CC used to illustrate the invention.  
 XX  
 SQ Sequence 628 AA;

Query Match 96.3%; Score 1293; DB 8; Length 628;  
 Best Local Similarity 95.7%; Pred. No. 6.5e-133;  
 Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 YEFLETIGKGTGKVKKARESSGRLVAIKSIRKDKIKDEQDLHRRREIEMSSLNHPHI 60  
 DB 53 YEFLETIGKGTGKVKKARESSGRLVAIKSIRKDKIKDEQDLHRRREIEMSSLNHPHI 112  
 QY 61 IAIHEVFENSKIVIVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 120  
 DB 113 IAIHEVFENSKIVIVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONGIV 172  
 QY 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPVVGPEVD 180  
 DB 173 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPVVGPEVD 232  
 QY 181 SWSGLVLLYLIVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDAGLIRLLMVNPTERRA 240  
 DB 233 SWSGLVLLYLIVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDAGLIRLLMVNPTERRA 292  
 QY 241 TLEDVASHWVWVW 253  
 DB 293 TLEDVASHWVWVW 305

RESULT 10  
 ADL14161  
 ID ADL14161 standard; protein; 628 AA.  
 XX  
 AC ADL14161;  
 DT 17-JUN-2004 (first entry)  
 XX  
 DE Novel human gene 3700 encoded protein.  
 XX  
 KW cytotostatic; cardiant; hypotensive; antianginal; osteopathic;  
 KW antiarthritic; antirheumatic; neuroprotective; antiinflammatory;  
 KW antipsoriatic; antiaesthatic; cardiovascular; virucide, analgesic; CNS;  
 KW angiogenesis inhibitor; angiogenesis stimulator; cerebroprotective;  
 KW nephrotropic; antithyroid; dermatological; immunomodulator;  
 KW cell proliferation disorder; cell differentiation disorder;  
 KW kidney disorder; renal disorder; lung disorder; ovarian disorder;  
 KW prostate disorder; cervical disorder; spleen disorder; thymus disorder;  
 KW thyroid disorder; testes disorder; haematopoietic disorder;  
 KW pancreatic disorder; skeletal muscle disorder; skin disorder;  
 KW dermal disorder; bone metabolism disorder; immune disorder;  
 KW inflammatory disorder; cardiovascular disorder;  
 KW endothelial cell disorder; liver disorder; viral disease; pain disorder;  
 KW metabolic disorder; neurological disorder;  
 KW central nervous system disorder; erythroid disorder;  
 KW blood vessel disorder; angiogenic disorder; cancer; heart failure;  
 KW hypertension; angina; osteoarthritis; rheumatoid arthritis;



KW multiple sclerosis; Crohn's disease; psoriasis; asthma;  
 KW cell proliferation; cell differentiation; cell growth; cell division;  
 KW human.  
 XX  
 OS Homo sapiens.  
 XX US2004058355-A1.  
 XX 25-MAR-2004.  
 XX 25-APR-2003; 2003US-004233543.  
 XX 30-SEP-1998; 98US-00163821.  
 XX 27-JAN-1999; 99US-0117580P.  
 PR 25-MAR-1999; 99US-00276400.  
 PR 30-JUL-1999; 99US-00365162.  
 PR 09-SEP-1999; 99US-00392189.  
 PR 05-OCT-1999; 99US-00412210.  
 PR 23-NOV-1999; 99US-00448076.  
 PR 29-FEB-2000; 2000US-0186061P.  
 PR 28-APR-2000; 2000US-0200688P.  
 PR 19-MAY-2000; 2000US-0205447P.  
 PR 30-JUN-2000; 2000US-00608921.  
 PR 31-JUL-2000; 2000US-0221925P.  
 PR 25-SEP-2000; 2000US-0234922P.  
 PR 25-SEP-2000; 2000US-0235035P.  
 PR 08-NOV-2000; 2000US-0246669P.  
 PR 09-NOV-2000; 2000US-00711216.  
 PR 14-NOV-2000; 2000US-0248335P.  
 PR 15-NOV-2000; 2000US-0248893P.  
 PR 22-DEC-2000; 2000US-0257511P.  
 PR 05-JAN-2001; 2001US-0260166P.  
 PR 28-FEB-2001; 2001US-00797039.  
 PR 27-APR-2001; 2001US-00845044.  
 PR 20-JUL-2001; 2001US-00909743.  
 PR 31-JUL-2001; 2001US-00920346.  
 PR 13-AUG-2001; 2001US-00928531.  
 PR 14-AUG-2001; 2001US-00929218.  
 PR 15-AUG-2001; 2001US-0312539P.  
 PR 25-SEP-2001; 2001US-00963159.  
 PR 08-NOV-2001; 2001US-0008016.  
 PR 13-NOV-2001; 2001US-00012055.  
 PR 15-NOV-2001; 2001US-00003690.  
 PR 30-JAN-2002; 2002US-00060763.  
 PR 25-MAR-2002; 2002US-00105989.  
 PR 12-APR-2002; 2002US-00121911.  
 PR 12-AUG-2002; 2002US-00217168.  
 PR 22-OCT-2002; 2002US-00278036.  
 PR 02-JAN-2003; 2003US-00336489.  
 PR 03-JAN-2003; 2003US-00336153.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX  
 XX Kapeller-Libermann R, Hunter JJ, Meyers RE, Rudolph-Owen LA;  
 PI Curtis RAJ, Olant PJ, Tsai F, Galvin KM, Chun M, Williamson MJ;  
 PI Silos-Santiago I, Bandaru R;  
 XX  
 XX MPI: 2004-268788/25.  
 DR N-PSDB; ADL14160, ADL14162.  
 DR  
 XX New 21910, 56634, 55053, 2504, 15977, 14760, 25501, 17903, 3700, 21529,  
 PT 26176, 26343, 56638, 18610, 33217, 21967, h1983, m1983, 38555 or 593  
 PT nucleic acid molecules and proteins, useful for treating, e.g. cancer,  
 PT heart failure and angina.  
 XX  
 PS Claim 4; SEQ ID NO 44; 139pp; English.  
 XX  
 CC The invention describes an isolated 21910, 56634, 55053, 2504, 15977,  
 CC 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217,  
 CC 21967, h1983, m1983, 38555 or 593 nucleic acid molecule (I) comprising  
 CC any one of 40 nucleotide sequences (i). The nucleic acid molecules and  
 CC polypeptides are useful for diagnosing and treating a subject having a  
 CC disorder, or a subject at risk of developing a disorder, which is

CC associated with aberrant 21910, 56634, 55053, 2504, 15977, 14760, 25501,  
 CC 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217, 21967, h1983,  
 CC m1983, 38555 or 593 activity, such as cellular proliferative and/or  
 CC differentiative disorders, brain disorders, platelet disorders, breast  
 CC disorders, colon disorders, kidney (renal) disorders, lung disorders,  
 CC ovarian disorders, prostate disorders, cervical disorders, spleen  
 CC disorders, thymus disorders, thyroid disorders, testes disorders,  
 CC hematopoietic disorders, pancreatic disorders, skeletal muscle disorders,  
 CC skin (dermal) disorders, disorders associated with bone metabolism,  
 CC immune, e.g. inflammatory disorders, cardiovascular disorders,  
 CC endothelial cell disorders, liver disorders, viral diseases, pain  
 CC disorders, metabolic disorders, neurological or central nervous system  
 CC disorders, erythroid disorders, blood vessel disorders or angiogenic  
 CC disorders (all claimed), e.g. cancer, heart failure, hypertension,  
 CC angina, osteoarthritis, rheumatoid arthritis, multiple sclerosis, Crohn's  
 CC disease, psoriasis, or asthma. The nucleic acid molecules and  
 CC polypeptides are also useful as modulating agents in regulating a variety  
 CC of cellular processes, e.g. cell proliferation, differentiation, growth and  
 CC division. This is the amino acid sequence of a novel human protein of the  
 CC invention. Note: The sequences given in the specification are also  
 CC available in electronic format from  
 CC ftp.seqdata.uspto.gov/sequence.html?DocID=20040058355.  
 XX  
 SQ Sequence 628 AA;  
 Qy 1 YFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEQDLHRIREIEMSSLNHPHI 60  
 Db 53 YFLETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEQDLHRIREIEMSSLNHPHI 112  
 Qy 61 IATHEVFENSISKIVIMEYASRGDLVDYISERPLSERDARHPROIVSALHYCHONGIV 120  
 Db 113 IATHEVFENSISKIVIMEYASRGDLVDYISERQQLSEREARHPROIVSALHYCHONRV 172  
 Qy 121 HRDLKLENIILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPIVNGKPYVGPEVD 180  
 Db 173 HRDLKLENIILLDANGNIKIADFGLSNLYHOGKFLQTFCGSPLYASPIVNGKPYTGPEVD 232  
 Qy 181 SWSLGVLLYILVHGTWPFDDQDHKTLVKQISNGAYREPPKPSDACGLIRLLMVPTRRA 240  
 Db 233 SWSLGVLLYILVHGTWPFDDQDHKTLVKQISNGAYREPPKPSDACGLIRLLMVPTRRA 292  
 Qy 241 TLEDVASHHWVNW 253  
 Db 293 TLEDVASHHWVNW 305  
 RESULT 11  
 ADO20172  
 ID ADO20172 standard; protein; 628 AA.  
 XX  
 AC ADO20172;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human PRO polypeptide #540.  
 XX  
 KW Human; PRO; immune related disorder; systemic lupus erythematosus;  
 KW rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;  
 KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;  
 KW autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;  
 KW diabetes mellitus; renal disease; demyelinating disease;  
 KW central nervous system; peripheral nervous system;  
 KW demyelinating polyneuropathy; Guillain-Barre syndrome;  
 KW chronic inflammatory demyelinating polyneuropathy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004043361-A2.  
 XX

PD 27-MAY-2004.  
 XX 06-NOV-2003; 2003WO-US035268.  
 XX 08-NOV-2002; 2002US-0425235P.  
 XX (GETH ) GENENTECH INC.  
 XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;  
 PI Wood WI, Wu TD;  
 XX WPI; 2004-420067/39.  
 DR N-PSDB; ADO20171.  
 XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for  
 PT treating an immune related disorder such as systemic lupus erythematosus,  
 PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or  
 PT spondyloarthritis.  
 XX Claim 7; SEQ ID NO 1080; 1731pp; English.  
 XX The invention relates to human PRO polypeptides and the polynucleotides  
 CC encoding them. The polypeptides and polynucleotides are useful for  
 CC treating and diagnosing immune related disorders in mammals. The immune  
 CC related disorders include systemic lupus erythematosus, rheumatoid  
 CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic  
 CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune  
 CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes  
 CC mellitus, immune-mediated renal disease, demyelinating diseases of the  
 CC central or peripheral nervous system, demyelinating polyneuropathy,  
 CC Guillain-Barre syndrome and chronic inflammatory demyelinating  
 CC polyneuropathy. This sequence represents a human PRO polypeptide of the  
 CC invention.  
 XX Sequence 628 AA;  
 SQ

Query Match 96.3%; Score 1293; DB 8; Length 628;  
 Best Local Similarity 95.7%; Pred. No. 6.5e-133;  
 Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 YEFLTLGKGYGKVKKARESSGRLVAIKSRKDKIKDEQDLHIREIEMSLNPHI 60  
 DB 53 YEFLTLGKGYGKVKKARESSGRLVAIKSRKDKIKDEQDLHIREIEMSLNPHI 112  
 QY 61 IAIHEVPENSSKIVIVMEYASRGDLYISERPLSERDARHFRQIVSAHYCHONGIV 120  
 DB 113 IAIHEVPENSSKIVIVMEYASRGDLYISERQQLSREARHFRQIVSAHYCHONGIV 172  
 QY 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTFGSPLYASPEIVNGKPYVGPVD 180  
 DB 173 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTFGSPLYASPEIVNGKPYTGPVD 232  
 QY 181 SWSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRLLMNVNPTTRA 240  
 DB 233 SWSLGVLLYILVHGTMPFDGHDHKLILVKQISNGAYREPPKPSDACGLIRLLMNVNPTTRA 292  
 QY 241 TLEDVASHWVWVW 253  
 DB 293 TLEDVASHWVWVW 305

RESULT 12  
 ADJ96620  
 ID ADJ96620 standard; protein; 672 AA.  
 XX  
 AC ADJ96620;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX Human calcium/calmodulin-dependent protein kinase NuaK2 protein SeqID 77.  
 DE kinase; human; tyrosine protein kinase; serine/threonine protein kinase;  
 KW PK; STK; Gene therapy; cancer; immune-related disease;  
 KW

KW cardiovascular disease; brain; neuronal associated disease; metabolic;  
 KW inflammatory disorder; cytostatic; neuroprotective; immunomodulator;  
 KW antiinflammatory; enzyme; calcium/calmodulin-dependent protein kinase;  
 KW NuaK2.  
 XX Homo sapiens.  
 OS 72.  
 XX WO2004006838-A2.  
 PN 22-JAN-2004.  
 XX 15-JUL-2003; 2003WO-US021730.  
 PD 15-JUL-2002; 2002US-0395632P.  
 XX (SUGE-) SUGEN INC.  
 PA Whyte D, Manning G, Caenepeel S;  
 PI WPI; 2004-122753/12.  
 DR N-PSDB; ADJ96554.  
 XX New nucleic acid molecule encoding a kinase polypeptide, useful for  
 PT preparing a composition for treating diseases or disorders, e.g., cancer,  
 PT or neurological, immunological or inflammatory disorders.  
 XX Claim 1; SEQ ID NO 77; 366pp; English.  
 XX This invention relates to a novel isolated, enriched or purified nucleic  
 CC acid molecule that encodes a kinase polypeptide. Specifically, it relates  
 CC to human tyrosine and serine/threonine protein kinases (PTK's and STK's),  
 CC as well as protein kinase-like enzymes. The present invention describes  
 CC screening methods to identify agonists, antagonists and antibodies that  
 CC can be used to modulate the activity or function of the mammalian kinase  
 CC enzymes. As such, these compositions can be used for gene therapy  
 CC purposes to treat diseases or disorders including cancer, immune-related  
 CC diseases, cardiovascular disease, brain or neuronal associated disease,  
 CC metabolic and inflammatory disorders. Accordingly, they exhibit  
 CC cytostatic, neuroprotective, immunomodulator and antiinflammatory  
 CC activities. This polypeptide sequence is a human kinase protein sequence  
 CC of the invention.  
 XX Sequence 672 AA;  
 SQ

Query Match 96.3%; Score 1293; DB 8; Length 672;  
 Best Local Similarity 95.7%; Pred. No. 7.2e-133;  
 Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 YEFLTLGKGYGKVKKARESSGRLVAIKSRKDKIKDEQDLHIREIEMSLNPHI 60  
 DB 97 YEFLTLGKGYGKVKKARESSGRLVAIKSRKDKIKDEQDLHIREIEMSLNPHI 156  
 QY 61 IAIHEVPENSSKIVIVMEYASRGDLYISERPLSERDARHFRQIVSAHYCHONGIV 120  
 DB 157 IAIHEVPENSSKIVIVMEYASRGDLYISERQQLSREARHFRQIVSAHYCHONGIV 216  
 QY 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTFGSPLYASPEIVNGKPYVGPVD 180  
 DB 217 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTFGSPLYASPEIVNGKPYTGPVD 276  
 QY 181 SWSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRLLMNVNPTTRA 240  
 DB 277 SWSLGVLLYILVHGTMPFDGHDHKLILVKQISNGAYREPPKPSDACGLIRLLMNVNPTTRA 336  
 QY 241 TLEDVASHWVWVW 253  
 DB 337 TLEDVASHWVWVW 349

RESULT 13  
 AAE16266  
 ID AAE16266 standard; protein; 594 AA.

XX AAE16266;  
XX 26-MAR-2002 (first entry)  
XX Human kinase PKIN-12 protein.  
XX Human; kinase; PKIN-12; cancer; leukaemia; adenocarcinoma; osteoporosis;  
XX immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;  
XX Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;  
XX allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;  
XX autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis;  
XX Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;  
XX rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;  
XX hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;  
XX cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;  
XX congestive heart failure; ischaemic heart disease; lung tumour; gout;  
XX fatty liver; Niemann-Pick's disease; gene therapy.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
XX Domain 19..269  
XX /note= "Eukaryotic protein kinase domain"  
XX Domain 22..260  
XX /label= Protein\_kinase\_domain  
XX Domain 23..260  
XX /label= Protein\_kinase\_domain  
XX Domain 24..260  
XX /label= Protein\_kinase\_domain  
XX WO200196547-A2.  
XX 20-DEC-2001.  
XX 14-JUN-2001; 2001WO-US019444.  
XX 15-JUN-2000; 2000US-0212073P.  
XX 23-JUN-2000; 2000US-0213467P.  
XX 30-JUN-2000; 2000US-0215651P.  
XX 07-JUL-2000; 2000US-0216605P.  
XX 13-JUL-2000; 2000US-0218372P.  
XX 25-AUG-2000; 2000US-0228056P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;  
XX Gandhi AR, Tribouley CM, Wallia NK, Yao MG, Lu DAM, Greenwald SR;  
XX Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang YT;  
XX Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;  
XX Lo TP, Khan F, Recipon SA, Azimzai V, Pollicky JL, Ding L;  
XX Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;  
XX WPI; 2002-090207/12.  
XX N-PSDB; AAD26459.  
XX New polypeptides, useful for diagnosing, treating or preventing disorders  
XX of growth and development, cardiovascular and lipid, and diseases such as  
XX cancer, comprise human kinase polypeptides.  
XX Claim 1; Page 152-153; 197pp; English.  
XX The invention relates to human kinase PKIN proteins and their  
XX corresponding cDNAs. A composition containing PKIN agonist is useful for  
XX treating a disease or condition associated with decreased expression of  
XX PKIN and a composition comprising PKIN antagonist is useful for treating  
XX a disease or condition associated with overexpression of PKIN. The  
XX disorders include cancer (leukaemia, adenocarcinoma, lymphoma, melanoma,  
XX myeloma, sarcoma, teratocarcinoma, Hodgkin's disease); immune disorder  
XX (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,  
XX atherosclerosis, anaemia, allergies, adult respiratory distress syndrome,  
XX autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes  
XX mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,  
XX osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,  
XX rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,  
XX bacterial, parasitic, fungal, viral, protozoal and helminthic infections)  
XX growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,  
XX Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); cardio  
XX vascular disease (arteriovenous fistula, hypertension, vasculitis,  
XX aneurysms, congestive heart failure, angina pectoris, myocarditis,  
XX ischaemic heart disease, chronic bronchitis, lung tumours); lipid  
XX disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,  
XX hypocholesterolaemia, obesity). PKIN DNA is useful for assessing toxicity  
XX of a test compound and in gene therapy. The present sequence is human  
XX PKIN-12 protein  
XX SQ Sequence 594 AA;  
XX Query Match 95.8%; Score 1286; DB 5; Length 594;  
XX Best Local Similarity 95.3%; Pred. No. 3.5e-132;  
XX Matches 241; Conservative 6; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 YEFLETLGKTYGKVKKARESSGRLVAIKSIRKDKIKDEQDLHRIREIEIMSSLNHPHI 60  
Db 19 YEFLETLGKTYGKVKKARESSGRLVAIKSIRKDKIKDEQDLHRIREIEIMSSLNHPHI 78  
Qy 61 IAIHEVPENSKIVIVMEYASRGDLVDYISERPLSRDARHPFQIVSAHYCHQNGIV 120  
Db 79 IAIHEVPENSKIVIVMEYASRGDLVDYISERQQLSREARHPFQIVSAHYCHQNRV 138  
Qy 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTECGSPLYASPEIVNGKPYVGPEVD 180  
Db 139 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTECGSPLYASPEIVNGKPYGTGPEVD 198  
Qy 181 SWSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRWLLMVPTRRA 240  
Db 199 SWSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRWLLMVPTRRA 258  
Qy 241 TLEDVASHWVWVW 253  
Db 259 TLEDVASHWVWVW 271  
RESULT 14  
ABP96085  
ID ABP96085 standard; protein; 611 AA.  
XX AC ABP96085;  
XX 07-MAY-2003 (first entry)  
XX Human protein kinase SEQ ID NO:76.  
XX Human; protein kinase; enzyme; antidiabetic; antidiabetic;  
XX antidiabetic; antiparkinsonian; antimigraine; cardiant; cytostatic;  
XX immunosuppressive; vulnery; gene therapy; COPD; asthma; migraine;  
XX chronic obstructive pulmonary disease; non-insulin dependent diabetes;  
XX Parkinson's disease; myocardial infarction; inflammatory bowel disease;  
XX autoimmune disorder; allograft rejection; graft versus host disease;  
XX cancer; leukaemia; wound granulation.  
XX Homo sapiens.  
XX WO2003000901-A2.  
XX 03-JAN-2003.  
XX 24-JUN-2002; 2002WO-IB002358.  
XX 26-JUN-2001; 2001US-0301098P.  
XX 06-NOV-2001; 2001US-0332870P.  
XX (DECO-) DECODE GENETICS EHF.  
XX Martinez RAM, Sigurdson GT;  
XX

```

DR WPI; 2003-201429/19.
XX N-PSDB; ABZ77163.
XX
PT New protein kinase genes and polypeptides, useful for diagnosing diseases
PT associated with a protein kinase, or in gene therapy for treating e.g.
PT Parkinson's disease, migraine, myocardial infarction, allograft rejection
PT or cancers.
XX
XX Claim 9; Page 87; 258pp; English.
XX
CC ABZ77126 to ABZ77165 encode the human protein kinases given in ABP96048
CC to ABP96087. The protein kinases have antiasthmatic, antiinflammatory,
CC antidiabetic, antiparkinsonian, antimigraine, cardiac, cytostatic,
CC immunosuppressive and vulnerary activities, and can be used in gene
CC therapy. A protein kinase therapeutic agent from the present invention,
CC particularly a protein kinase gene agonist or antagonist, can be used for
CC treating a disease or condition associated with a protein kinase in an
CC individual. These diseases include chronic obstructive pulmonary diseases
CC (COPD), asthma, non-insulin dependent diabetes, Parkinson's disease,
CC migraine, myocardial infarction, inflammatory bowel disease, autoimmune
CC disorders (e.g. allograft rejection or graft vs. host disease), cancers
CC (e.g. leukaemia) or wound granulation
XX
XX Sequence 611 AA;
XX
Query Match 95.3%; Score 1279.5; DB 6; Length 611;
Best Local Similarity 93.1%; Pred. No. 1.9e-131;
Matches 242; Conservative 6; Mismatches 5; Indels 7; Gaps 1;
XX
QY 1 YFELETGKGTGKVKKARESSGRLVAIKSIRKDKIKDEQDLHRIREIEMSSLNHPHI 60
Db 53 YFELETGKGTGKVKKARESSGRLVAIKSIRKDKIKDEQDLHRIREIEMSSLNHPHI 112
QY 61 IATHEVFENSCKIVIVMEYASRGDLVDYISERPLSERDARHFRQIVSAHYCHQNGIV 120
Db 113 IATHEVFENSCKIVIVMEYASRGDLVDYISERQQLSREARHFRQIVSAHYCHQNRVV 172
QY 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTFCGSPLYASPIVNGKPVVGPVD 180
Db 173 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTFCGSPLYASPIVNGKPYTGPEVD 232
QY 181 SWSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYREPPKPS-----DACGLIRWLLM 233
Db 233 SWSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYREPPKPSAPAFCLPDACGLIRWLLM 292
QY 234 VNPTRATLEDVASHWVWVNW 253
Db 293 VNPTRATLEDVASHWVWVNW 312
XX
RESULT 15
ID AAB65632 standard; protein; 629 AA.
XX
AC AAB65632;
XX
XX 27-MAR-2001 (first entry)
XX
DE Novel protein kinase, SEQ ID NO: 159.
XX
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
XX
OS Homo sapiens.
XX
PN WO200073469-A2.
XX
PD 07-DEC-2000.
XX
XX

```

```

PF 26-MAY-2000; 2000WO-US014842.
XX
PR 28-MAY-1999; 99US-0136503P.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX
XX WPI; 2001-032161/04.
DR N-PSDB; AAF44659.
XX
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
XX treating immune-related diseases and disorders, cardiovascular disease,
XX neurodegenerative diseases and/or cancers.
XX
PS Claim 10; Fig 1; 310pp; English.
XX
XX The present sequence is a novel protein kinase. The novel protein kinases
XX and the nucleic acids that encode them may be used in the treatment and
XX diagnosis of diseases associated with inappropriate kinase expression
XX such as immune-related diseases and disorders, cardiovascular disease,
XX neurodegenerative diseases and/or cancers. The nucleic acids and
XX complementary sequences may also be used as DNA probes in diagnostic
XX assays. The kinase polypeptides may be used as antigens in the production
XX of antibodies of kinase expression and activity. Anti-kinase antibodies
XX and kinase antagonists may also be used to down regulate kinase
XX expression and activity. Diseases related to kinase expression and
XX activity include rheumatoid arthritis, atherosclerosis, autoimmune
XX disorders, complications of organ transplantation, myocardial infarction,
XX immune disorders, cardiomyopathies, strokes, renal failure, oxidative-
XX stress related disorders, chronic inflammatory bowel disease, chronic
XX inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,
XX psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive
XX disorders
XX
XX Sequence 629 AA;
XX
Query Match 94.5%; Score 1268.5; DB 4; Length 629;
Best Local Similarity 94.5%; Pred. No. 3.3e-130;
Matches 240; Conservative 6; Mismatches 7; Indels 1; Gaps 1;
XX
QY 1 YFELETGKGTGKVKKARESSGRLVAIKSIRKDKIKDEQDLHRIREIEMSSLNHPHI 60
Db 53 YFELETGKGTGKVKKARESSGRLVAIKSIRKDKIKDEQDLHRIREIEMSSLNHPHI 112
QY 61 IATHEVFENSCKIVIVMEYASRGDLVDYISERPLSERDARHFRQIVSAHYCHQNGIV 120
Db 113 IATHEVFENSCKIVIVMEYASRGDLVDYISERQQLSREARHFRQIVSAHYCHQNRVV 172
QY 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTFCGSPLYASPIVNGKPVVGPVD 180
Db 173 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTFCGSPLYASPIVNGKPYTGPEVD 232
QY 181 SWSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDAC-GLIRWLLMVNPTRR 239
Db 233 SWSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDCLAGLIRWLLMVNPTRR 292
QY 240 ATLEDVASHWVWVNW 253
Db 293 ATLEDVASHWVWVNW 306
XX
Search completed: January 22, 2005, 06:38:27
Job time : 64.8224 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 06:43:09 ; Search time 58.6708 Seconds  
(without alignments)  
1557.950 Million cell updates/sec

Title: US-09-980-464-11\_COPY\_57\_309

Perfect score: 1342  
Sequence: 1 YEFLETGKTYGKVKKARE.....VNPRTRATLEDVASHMVVNW 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1342	100.0	631	14	US-10-355-975-11
2	1325	98.7	251	16	US-10-343-514-103
3	1313	97.8	630	16	US-10-343-514-41
4	1296	96.6	251	16	US-10-343-514-50
5	1293	96.3	628	9	US-09-963-159-2
6	1293	96.3	628	15	US-10-423-543-44
7	1293	96.3	628	17	US-10-370-715B-640
8	1293	96.3	672	17	US-10-618-941-77
9	1286	95.8	594	15	US-10-311-034-12
10	1276	95.1	251	16	US-10-343-514-101
11	1273	94.9	640	16	US-10-322-281-23
12	1230	91.7	616	16	US-10-322-281-26
13	1128.5	84.1	661	9	US-09-780-949-2
Sequence 11, Appl					
Sequence 103, Appl					
Sequence 41, Appl					
Sequence 50, Appl					
Sequence 2, Appl					
Sequence 44, Appl					
Sequence 640, Appl					
Sequence 77, Appl					
Sequence 12, Appl					
Sequence 101, Appl					
Sequence 26, Appl					
Sequence 12, Appl					

14	1128.5	84.1	661	9	US-09-780-949-6	Sequence 6, Appl
15	1128.5	84.1	661	14	US-10-354-358-82	Sequence 82, Appl
16	1128.5	84.1	661	17	US-10-723-860-2553	Sequence 2553, Appl
17	1111.5	82.8	252	16	US-10-343-514-102	Sequence 102, Appl
18	814	60.7	530	9	US-09-836-392-20	Sequence 20, Appl
19	785.5	58.5	246	10	US-09-898-837A-29	Sequence 29, Appl
20	665.5	49.6	508	15	US-10-016-248-24	Sequence 24, Appl
21	665.5	49.6	639	15	US-10-016-248-71	Sequence 71, Appl
22	665.5	49.6	639	15	US-10-016-248-72	Sequence 72, Appl
23	665.5	49.6	688	15	US-10-276-645-8	Sequence 8, Appl
24	665.5	49.6	703	15	US-10-016-248-70	Sequence 70, Appl
25	665.5	49.6	732	9	US-09-835-081-2	Sequence 2, Appl
26	665.5	49.6	752	15	US-10-258-106-16	Sequence 16, Appl
27	665.5	49.6	752	15	US-10-276-645-7	Sequence 7, Appl
28	665.5	49.6	769	15	US-10-363-616-403	Sequence 403, Appl
29	665.5	49.6	825	15	US-10-425-114-54516	Sequence 54516, A
30	661.5	49.3	752	17	US-10-618-941-79	Sequence 79, Appl
31	660.5	49.2	639	15	US-10-016-248-73	Sequence 73, Appl
32	660.5	49.2	660	15	US-10-276-645-6	Sequence 6, Appl
33	660.5	49.2	688	14	US-10-161-565-28	Sequence 28, Appl
34	660.5	49.2	688	14	US-10-161-565-29	Sequence 29, Appl
35	660.5	49.2	724	15	US-10-276-645-5	Sequence 5, Appl
36	651.5	48.5	744	9	US-09-835-081-4	Sequence 4, Appl
37	649.5	48.4	729	14	US-10-142-356-11	Sequence 11, Appl
38	649.5	48.4	729	14	US-10-195-101-33	Sequence 33, Appl
39	649.5	48.4	729	14	US-10-161-565-26	Sequence 26, Appl
40	649.5	48.4	744	9	US-09-919-585-3	Sequence 3, Appl
41	647.5	48.2	691	9	US-09-919-585-6	Sequence 6, Appl
42	647.5	48.2	722	14	US-10-274-194-4	Sequence 4, Appl
43	647.5	48.2	722	16	US-10-760-407-4	Sequence 4, Appl
44	647.5	48.2	724	9	US-09-919-585-9	Sequence 9, Appl
45	647.5	48.2	745	14	US-10-195-101-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1

US-10-355-975-11  
; Sequence 11, Application US/10355975  
; Publication No. US20030162277A1  
; GENERAL INFORMATION:  
; APPLICANT: Immunex Corporation  
; APPLICANT: Bird, Timothy A.  
; APPLICANT: Virca, G. Duke  
; APPLICANT: Martin, Unja  
; APPLICANT: Anderson, Dirk M.  
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES  
; FILE REFERENCE: 2923-A  
; CURRENT FILING DATE: 2003-01-30  
; PRIOR APPLICATION NUMBER: US/10/355,975  
; PRIOR FILING DATE: 2003-01-30  
; PRIOR APPLICATION NUMBER: US/09/579,664B  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 11  
; LENGTH: 631  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-355-975-11

Query Match 100.0%; Score 1342; DB 14; Length 631;  
Best Local Similarity 100.0%; Pred. No. 9.4e-100;  
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	YEFLETGKTYGKVKKARESSGRLVAIKSRDKIKDEODLLHIRREIETMSLNPHI	60
Db	57	YEFLETGKTYGKVKKARESSGRLVAIKSRDKIKDEODLLHIRREIETMSLNPHI	116
Qy	61	IATHEVFENSKIVIVMEYASRGDLXYISERPLSRDARHFFRQIVSALHYCHONGIV	120
Db	117	IATHEVFENSKIVIVMEYASRGDLXYISERPLSRDARHFFRQIVSALHYCHONGIV	176

Qy 121 HRDLKENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 180  
Db 177 HRDLKENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 236  
Qy 181 SWSLGVLLYILVHGTMPDGDHKTIVKQISNGAYREPPKPSDACGLIRLLMWNPTTRA 240  
Db 237 SWSLGVLLYILVHGTMPDGDHKTIVKQISNGAYREPPKPSDACGLIRLLMWNPTTRA 296  
Qy 241 TLEDVASHWVNW 253  
Db 297 TLEDVASHWVNW 309

## RESULT 2

US-10-343-514-103  
; Sequence 103, Application US/10343514  
; Publication No. US20040132025A1  
; GENERAL INFORMATION:  
; APPLICANT: DRUCKER, Daniel J.  
; APPLICANT: ROSEN, Cheryl F.  
; APPLICANT: LEBEVRE, Diana L.  
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK  
; FILE REFERENCE: DPA-DRUC2/PCT  
; CURRENT APPLICATION NUMBER: US/10/343.514  
; CURRENT FILING DATE: 2003-01-31  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR APPLICATION NUMBER: PCT/CA01/01109  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/222,650  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/274,613  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: CA 2,340,780  
; PRIOR FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 109  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 103  
; LENGTH: 251  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-343-514-103

Query Match 98.7%; Score 1325; DB 16; Length 251;  
Best Local Similarity 100.0%; Pred. No. 8e-99;  
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YEFLETGKGTGKGVKKARESSGRLVAIKSRKDKIKDEQDLHHRREIETMSSLNHPHI 60  
Db 1 YEFLETGKGTGKGVKKARESSGRLVAIKSRKDKIKDEQDLHHRREIETMSSLNHPHI 60  
Qy 61 IAHEVPENSSKIVVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHQNGIV 120  
Db 61 IAHEVPENSSKIVVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHQNGIV 120  
Qy 121 HRDLKENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 180  
Db 121 HRDLKENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 180  
Qy 181 SWSLGVLLYILVHGTMPDGDHKTIVKQISNGAYREPPKPSDACGLIRLLMWNPTTRA 240  
Db 181 SWSLGVLLYILVHGTMPDGDHKTIVKQISNGAYREPPKPSDACGLIRLLMWNPTTRA 240  
Qy 241 TLEDVASHWV 251  
Db 241 TLEDVASHWV 251

## RESULT 3

US-10-343-514-41  
; Sequence 41, Application US/10343514  
; Publication No. US20040132025A1  
; GENERAL INFORMATION:  
; APPLICANT: DRUCKER, Daniel J.  
; APPLICANT: ROSEN, Cheryl F.

; APPLICANT: LEBEVRE, Diana L.  
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK  
; FILE REFERENCE: DPA-DRUC2/PCT  
; CURRENT APPLICATION NUMBER: US/10/343.514  
; CURRENT FILING DATE: 2003-01-31  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR APPLICATION NUMBER: PCT/CA01/01109  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/222,650  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/274,613  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: CA 2,340,780  
; PRIOR FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 109  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 41  
; LENGTH: 630  
; TYPE: PRT  
; ORGANISM: RAT  
US-10-343-514-41

Query Match 97.8%; Score 1313; DB 16; Length 630;  
Best Local Similarity 98.0%; Pred. No. 2e-97;  
Matches 248; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YEFLETGKGTGKGVKKARESSGRLVAIKSRKDKIKDEQDLHHRREIETMSSLNHPHI 60  
Db 57 YEFLETGKGTGKGVKKARESSGRLVAIKSRKDKIKDEQDLHHRREIETMSSLNHPHI 116  
Qy 61 IAHEVPENSSKIVVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHQNGIV 120  
Db 117 IAHEVPENSSKIVVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHQNGIV 176  
Qy 121 HRDLKENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 180  
Db 177 HRDLKENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 236  
Qy 181 SWSLGVLLYILVHGTMPDGDHKTIVKQISNGAYREPPKPSDACGLIRLLMWNPTTRA 240  
Db 237 SWSLGVLLYILVHGTMPDGDHKTIVKQISNGAYREPPKPSDACGLIRLLMWNPTTRA 296  
Qy 241 TLEDVASHWVNW 253  
Db 297 TLEDVASHWVNW 309

## RESULT 4

US-10-343-514-50  
; Sequence 50, Application US/10343514  
; Publication No. US20040132025A1  
; GENERAL INFORMATION:  
; APPLICANT: DRUCKER, Daniel J.  
; APPLICANT: ROSEN, Cheryl F.  
; APPLICANT: LEBEVRE, Diana L.  
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK  
; FILE REFERENCE: DPA-DRUC2/PCT  
; CURRENT APPLICATION NUMBER: US/10/343.514  
; CURRENT FILING DATE: 2003-01-31  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR APPLICATION NUMBER: PCT/CA01/01109  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/222,650  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/274,613  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: CA 2,340,780  
; PRIOR FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 109  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 50  
; LENGTH: 251  
; TYPE: PRT  
; ORGANISM: RAT  
US-10-343-514-50

Query Match 96.6%; Score 1296; DB 16; Length 251;  
Best Local Similarity 98.0%; Pred. No. 1.7e-96;  
Matches 246; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YEFLETGKGTGKVKKARESSGRLVAIKSRKDKIKDEQDLLHRRREIEMSSLNHPHI 60  
Db |||||  
1 YEFLETGKGTGKVKKARESSGRLVAIKSRKDKIKDEQDLLHRRREIEMSSLNHPHI 60  
Qy 61 IAIHEVFENSSKIVIMVEYASRGDLVDYISERPLSERDARHFRQIVSALHYCHONGIV 120  
Db |||||  
61 IAIHEVFENSSKIVIMVEYASRGDLVDYISERPLSERDARHFRQIVSALHYCHONGIV 120  
Qy 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGFQTFCGSPLYASPEIVNGKPYVGPEVD 180  
Db |||||  
121 HRDLKLENILLDANGNIKIADFGLSNLYHKGFQTFCGSPLYASPEIVNGKPYVGPEVD 180  
Qy 181 SWSGLVLLYILVHGTMPDGDQDKTLVKQISNGAYREPPKPSDACGLIRWLLMNPTRRA 240  
Db |||||  
181 SWSGLVLLYILVHGTMPDGDQDKTLVKQISNGAYREPPKPSDACGLIRWLLMNPTRRA 240  
Qy 241 TLEDVASHWV 251  
Db |||||  
241 TLEDVASHWV 251

## RESULT 5

US-09-963-159-2  
; Sequence 2, Application US/09963159  
; Patent No. US2002007312A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtiss, Rory A.J.  
; APPLICANT: Galvin, Katherine M.  
; TITLE OF INVENTION: 3700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR  
; FILE REFERENCE: 10147-5001  
; CURRENT APPLICATION NUMBER: US/09/963,159  
; CURRENT FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 60/234,922  
; PRIOR FILING DATE: 2000-09-25  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 628  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-963-159-2

Query Match 96.3%; Score 1293; DB 9; Length 628;  
Best Local Similarity 95.7%; Pred. No. 8.4e-96;  
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 YEFLETGKGTGKVKKARESSGRLVAIKSRKDKIKDEQDLLHRRREIEMSSLNHPHI 60  
Db 53 YEFLETGKGTGKVKKARESSGRLVAIKSRKDKIKDEQDLLHRRREIEMSSLNHPHI 112  
Qy 61 IAIHEVFENSSKIVIMVEYASRGDLVDYISERPLSERDARHFRQIVSALHYCHONGIV 120  
Db 113 IAIHEVFENSSKIVIMVEYASRGDLVDYISERQQLSREARHFRQIVSAHYCHONRV 172  
Qy 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGFQTFCGSPLYASPEIVNGKPYVGPEVD 180  
Db 173 HRDLKLENILLDANGNIKIADFGLSNLYHQGFQTFCGSPLYASPEIVNGKPYTGPEVD 232  
Qy 181 SWSGLVLLYILVHGTMPDGDQDKTLVKQISNGAYREPPKPSDACGLIRWLLMNPTRRA 240  
Db 233 SWSGLVLLYILVHGTMPDGDQDKTLVKQISNGAYREPPKPSDACGLIRWLLMNPTRRA 292  
Qy 241 TLEDVASHWV 253  
Db 293 TLEDVASHWV 305

## RESULT 6

US-10-423-543-44  
; Sequence 44, Application US/10423543  
; Publication No. US20040058355A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Libermann, Rosana K.  
; APPLICANT: Hunter, John J.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Rudolph-Owen, Laura A.  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Tsai, Fong-Ying  
; APPLICANT: Galvin, Katherine M.  
; APPLICANT: Chun, Miyoung  
; APPLICANT: Williamson, Mark J.  
; APPLICANT: Silos-Santiago, Inmaculada  
; APPLICANT: Bandaru, Rajasekhar  
; TITLE OF INVENTION: NOVEL 21910, 56634, 2504, 15977,  
; TITLE OF INVENTION: 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638,  
; TITLE OF INVENTION: 18610, 33217, 21967, h1983, m1983, 38555 OR 593 MOLECULES  
; TITLE OF INVENTION: AND USES THEREFOR  
; FILE REFERENCE: MPI03-023OMNIM  
; CURRENT APPLICATION NUMBER: US/10/423,543  
; CURRENT FILING DATE: 2003-04-25  
; PRIOR APPLICATION NUMBER: US 10/278,036  
; PRIOR FILING DATE: 2002-10-22  
; PRIOR APPLICATION NUMBER: US 09/711,216  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: US 60/205,447  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 10/012,055  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/248,325  
; PRIOR FILING DATE: 2000-11-14  
; PRIOR APPLICATION NUMBER: US 10/003,690  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/248,893  
; PRIOR FILING DATE: 2000-11-15  
; PRIOR APPLICATION NUMBER: US 09/797,039  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/186,061  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 10/217,168  
; PRIOR FILING DATE: 2002-08-12  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44  
; LENGTH: 628  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-423-543-44

Query Match 96.3%; Score 1293; DB 15; Length 628;  
Best Local Similarity 95.7%; Pred. No. 8.4e-96;  
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 YEFLETGKGTGKVKKARESSGRLVAIKSRKDKIKDEQDLLHRRREIEMSSLNHPHI 60  
Db 53 YEFLETGKGTGKVKKARESSGRLVAIKSRKDKIKDEQDLLHRRREIEMSSLNHPHI 112  
Qy 61 IAIHEVFENSSKIVIMVEYASRGDLVDYISERPLSERDARHFRQIVSALHYCHONGIV 120  
Db 113 IAIHEVFENSSKIVIMVEYASRGDLVDYISERQQLSREARHFRQIVSAHYCHONRV 172  
Qy 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGFQTFCGSPLYASPEIVNGKPYVGPEVD 180  
Db 173 HRDLKLENILLDANGNIKIADFGLSNLYHQGFQTFCGSPLYASPEIVNGKPYTGPEVD 232  
Qy 181 SWSGLVLLYILVHGTMPDGDQDKTLVKQISNGAYREPPKPSDACGLIRWLLMNPTRRA 240  
Db 233 SWSGLVLLYILVHGTMPDGDQDKTLVKQISNGAYREPPKPSDACGLIRWLLMNPTRRA 292

```
Qy 241 TLEDVASHWVNW 253
Db 293 TLEDVASHWVNW 305

RESULT 7
US-10-370-715B-640
; Sequence 640, Application US/10370715B
; Publication No. US20040258678A1
; GENERAL INFORMATION:
;   Patin Docket Preview
;   APPLICANT: BODARY, SARAH C.
;   APPLICANT: CLARK, HILLARY
;   APPLICANT: BRISDELL, HUNTE
;   APPLICANT: JACKMAN, JANET
;   APPLICANT: SCHOENFELD, JILL R.
;   APPLICANT: WILLIAMS, P. MICKY
;   APPLICANT: WOOD, WILLIAM I.
;   APPLICANT: WU, THOMAS D.
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune
; FILE REFERENCE: P1948R1-US
; CURRENT APPLICATION NUMBER: US/10/370.715B
; CURRENT FILING DATE: 2003-02-21
; NUMBER OF SEQ ID NOS: 742
; SEQ ID NO 640
; LENGTH: 628
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-370-715B-640

Query Match          96.3%; Score 1293; DB 17; Length 628;
Best Local Similarity 95.7%; Pred. No. 8.4e-96;
Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 YEFLETGKGTGKGVKKARESSGRLVAIKSRKDKIKDEQDLHRRREIEIMSSLNHPHI 60
Db 53 YEFLETGKGTGKGVKKARESSGRLVAIKSRKDKIKDEQDLHRRREIEIMSSLNHPHI 112
Qy 61 IAIHEVFENSKIVIVMEYASRGDLVDYISERPLSRDARHPFROIIVSALHYCHONGIV 120
Db 113 IAIHEVFENSKIVIVMEYASRGDLVDYISERPLSRDARHPFROIIVSALHYCHONGIV 172
Qy 121 HRDLKLENILLDANGNIKIADFGLSNLYHGKFLQTFGSGPLYPASPEIVNGKPYVVGPEVD 180
Db 173 HRDLKLENILLDANGNIKIADFGLSNLYHGKFLQTFGSGPLYPASPEIVNGKPYVVGPEVD 232
Qy 181 SWSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRLLMVNPTTRA 240
Db 233 SWSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRLLMVNPTTRA 292
Qy 241 TLEDVASHWVNW 253
Db 293 TLEDVASHWVNW 305

RESULT 8
US-10-618-941-77
; Sequence 77, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
;   APPLICANT: WHYTE, DAVID
;   APPLICANT: MANNING, GERARD
;   APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618.941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 77

Qy 241 TLEDVASHWVNW 253
Db 337 TLEDVASHWVNW 349

RESULT 9
US-10-311-034-12
; Sequence 12, Application US/10311034
; Publication No. US20040023242A1
; GENERAL INFORMATION:
;   APPLICANT: INCYTE GENOMICS, INC.
;   APPLICANT: YUE, Henry
;   APPLICANT: LAL, Preeti
;   APPLICANT: BANDMAN, Olga
;   APPLICANT: BOROWSKY, Mark L.
;   APPLICANT: AU-YOUNG, Janice
;   APPLICANT: LU, Yan
;   APPLICANT: GANDHI, Ameenah R.
;   APPLICANT: TRIBOULEY, Catherine M.
;   APPLICANT: CHAWLA, Narinder K.
;   APPLICANT: YAO, Monique G.
;   APPLICANT: LU, Dyrung Aina M.
;   APPLICANT: GREENWALD, Sara R.
;   APPLICANT: RAMKUMAR, Jayalaxmi
;   APPLICANT: GRIFFIN, Jennifer A.
;   APPLICANT: KEARNEY, Liam
;   APPLICANT: BURFORD, Neil
;   APPLICANT: NGUYEN, Dannel B.
;   APPLICANT: TANG, Y. Tom
;   APPLICANT: BAUGHN, Mariah R.
;   APPLICANT: HE, Ann
;   APPLICANT: THORNTON, Michael
;   APPLICANT: HAFALIA, April
;   APPLICANT: ARVIZU, Chandra S.
;   APPLICANT: GURURAJAN, Rajagopal
;   APPLICANT: LO, Terence P.
;   APPLICANT: KHAH, Farrah A.
;   APPLICANT: RECIPON, Shirley A.
;   APPLICANT: AZIMZAI, Yalda
;   APPLICANT: POLICKY, Jennifer L.
;   APPLICANT: DING, Li
;   APPLICANT: GREYER, Megan
;   APPLICANT: ELLIOTT, Vicki S.
;   APPLICANT: THANGAVELU, Kavitha
;   APPLICANT: BATRA, Sajeew
;   APPLICANT: ISCON, Craig H.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0125 PCT
```



Query Match 95.1%; Score 1276; DB 16; Length 251;  
Best Local Similarity 95.6%; Pred. No. 7.2e-95;  
Matches 240; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

RESULT 12  
US-10-322-281-26  
; Sequence 26, Application US/10322281  
; Publication No. US20040126762A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro

;; TITLE OF INVENTION: Novel Compositions and Methods in Cancer

;; FILE REFERENCE: 529452001000  
;; CURRENT APPLICATION NUMBER: US/10/322,281  
;; CURRENT FILING DATE: 2002-12-17  
;; NUMBER OF SEQ ID NOS: 866  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 26  
;; LENGTH: 616  
;; TYPE: PRP  
;; ORGANISM: Homo sapiens  
US-10-322-281-26

Query Match 91.7%; Score 1230; DB 16; Length 616;  
Best Local Similarity 93.3%; Pred. No. 9.8e-91;  
Matches 236; Conservative 6; Mismatches 5; Indels 6; Gaps 3;  
Qy 1 YEFLETGKGTGYGVKKARESSGRVVAIKSIRKDKIKDEQDLHIRREIEMSSLNHPHI 60  
Db 53 YEFLETGKGTGYGVKKARESSGRVVAIKSIRKDKIKDEQDLHIRREIEMSSLNHPHI 112  
Qy 61 IAIHEVFENSKIVIMVEYASRGDLVDYISERPLSERDARHPFQIVSALHYCHONGIV 120  
Db 113 IAIHEVFENSKIVIMVEYASRGDLVDYISQ--QLSE--ARHFFRQIVSAVHYCHQNRV 168  
Qy 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEVD 180  
Db 169 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLY--DELVNGKPYVGPEVD 226  
Qy 181 SWSLGVLLYLIVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRWLLMVNPTRR 240  
Db 227 SWSLGVLLYLIVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRWLLMVNPTRR 286  
Qy 241 TLEDVASHWVWVNW 253  
Db 287 TLEDVASHWVWVNW 299

## RESULT 13

US-09-780-949-2  
;; Sequence 2, Application US/09780949  
;; Patent No. US20020006618A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Kapeller-Liberman, Rosana  
;; APPLICANT: Welch, Nadine S.  
;; APPLICANT: Galvin, Katherine M.  
;; TITLE OF INVENTION: Methods for Using 20893, a Human Protein  
;; FILE REFERENCE: 035800/209015  
;; CURRENT APPLICATION NUMBER: US/09/780,949  
;; CURRENT FILING DATE: 2001-02-09  
;; PRIOR APPLICATION NUMBER: US 60/181,690  
;; PRIOR FILING DATE: 2000-02-09  
;; NUMBER OF SEQ ID NOS: 6  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 2  
;; LENGTH: 661  
;; TYPE: PRP  
;; ORGANISM: H. sapiens  
US-09-780-949-2

Query Match 84.1%; Score 1128.5; DB 9; Length 661;  
Best Local Similarity 82.3%; Pred. No. 1.6e-82;  
Matches 209; Conservative 24; Mismatches 20; Indels 1; Gaps 1;  
Qy 1 YEFLETGKGTGYGVKKARESSGRVVAIKSIRKDKIKDEQDLHIRREIEMSSLNHPHI 59  
Db 55 YELQETLGKGTGYGVKKARESSGRVVAIKSIRKDKIKDEQDMVHIREIEMSSLNHPHI 114  
Qy 60 IAIHEVFENSKIVIMVEYASRGDLVDYISERPLSERDARHPFQIVSALHYCHONGI 119  
Db 115 IISIVYFENKDKIIVIMEYASKGELYDISSRRLSERTRHFFRQIVSAVHYCHNGV 174  
Qy 120 VHRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEV 179

Db 175 VHRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGRPGPEV 234  
Qy 180 DSWSLGVLLYLIVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRWLLMVNPTRR 239  
Db 235 DSWALGVLLYLIVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRWLLMVNPTRR 294  
Qy 240 ATLEDVASHWVWVNW 253  
Db 295 ATLEDVASHWVWVNW 308

## RESULT 14

US-09-780-949-6  
;; Sequence 6, Application US/09780949  
;; Patent No. US20020006618A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Kapeller-Liberman, Rosana  
;; APPLICANT: Welch, Nadine S.  
;; APPLICANT: Galvin, Katherine M.  
;; TITLE OF INVENTION: Methods for Using 20893, a Human Protein  
;; FILE REFERENCE: 035800/209015  
;; CURRENT APPLICATION NUMBER: US/09/780,949  
;; CURRENT FILING DATE: 2001-02-09  
;; PRIOR APPLICATION NUMBER: US 60/181,690  
;; PRIOR FILING DATE: 2000-02-09  
;; NUMBER OF SEQ ID NOS: 6  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 6  
;; LENGTH: 661  
;; TYPE: PRP  
;; ORGANISM: H. sapiens  
US-09-780-949-6

Query Match 84.1%; Score 1128.5; DB 9; Length 661;  
Best Local Similarity 82.3%; Pred. No. 1.6e-82;  
Matches 209; Conservative 24; Mismatches 20; Indels 1; Gaps 1;  
Qy 1 YEFLETGKGTGYGVKKARESSGRVVAIKSIRKDKIKDEQDLHIRREIEMSSLNHPHI 59  
Db 55 YELQETLGKGTGYGVKKARESSGRVVAIKSIRKDKIKDEQDMVHIREIEMSSLNHPHI 114  
Qy 60 IAIHEVFENSKIVIMVEYASRGDLVDYISERPLSERDARHPFQIVSALHYCHONGI 119  
Db 115 IISIVYFENKDKIIVIMEYASKGELYDISSRRLSERTRHFFRQIVSAVHYCHNGV 174  
Qy 120 VHRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEV 179  
Db 175 VHRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGRPGPEV 234  
Qy 180 DSWSLGVLLYLIVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRWLLMVNPTRR 239  
Db 235 DSWALGVLLYLIVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDACGLIRWLLMVNPTRR 294  
Qy 240 ATLEDVASHWVWVNW 253  
Db 295 ATLEDVASHWVWVNW 308

## RESULT 15

US-10-354-358-82  
;; Sequence 82, Application US/10354358  
;; Publication No. US20030157082A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Millennium Pharmaceuticals, Inc  
;; APPLICANT: Hunter, John Joseph  
;; APPLICANT: MacBeth, Kyle J.  
;; APPLICANT: Tsai, Fong-Ying  
;; APPLICANT: Lesoon, Andrea  
;; APPLICANT: Lightcap, Eric S.  
;; APPLICANT: Williamson, Mark  
;; APPLICANT: Rudolph-Owen, Laura A.





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 06:29:16 ; Search time 15.4548 Seconds  
(without alignments)  
1575.104 Million cell updates/sec

Title: US-09-980-464-11\_COPY\_57\_309  
Perfect score: 1342  
Sequence: 1 YEFLETGKTYGKVKARE.....VNPTTRATLEDVASHMWVN 253

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	813.5	60.6	1558	2 T29253	hypothetical prote
2	647.5	48.2	745	2 G01025	serine/threonine p
3	646.5	48.2	713	2 S27966	probable serine/th
4	644	48.0	512	2 T52633	serine/threonine-s
5	641	47.8	504	2 T10449	probable serine/th
6	638.5	47.6	1398	2 T13741	hypothetical prote
7	637	47.5	512	1 JC1446	serine/threonine-s
8	631	47.0	511	1 A56009	serine/threonine-s
9	627.5	46.8	774	2 T48609	probable serine/th
10	621.5	46.3	1192	2 T18611	probable serine/th
11	621.5	46.3	1246	2 G89287	protein H3923.1 f
12	621	46.3	472	2 B90100	SNF-related kinase
13	619.5	46.2	633	1 A26030	serine/threonine-s
14	618	46.1	504	2 T07415	probable serine/th
15	617.5	46.0	798	2 JC7500	gk protein - chic
16	616	45.9	562	2 T29858	hypothetical prote
17	614.5	45.8	602	2 S72513	FOG2 protein - yea
18	610	45.5	473	1 S59941	serine/threonine-s
19	604	45.0	552	1 S51025	[hydroxymethyl]glut
20	601	44.8	552	1 A53621	[hydroxymethyl]glut
21	599	44.6	512	2 T07788	probable serine/th
22	591.5	44.1	576	2 T41587	probable carbon ca
23	583.5	43.5	887	2 T20941	hypothetical prote
24	579.5	43.2	481	2 T49072	protein kinase - m
25	576.5	43.0	510	2 T04145	serine/threonine p
26	573	42.7	622	1 S44859	serine/threonine-s
27	567.5	42.3	502	1 A41361	serine/threonine-s
28	566.5	42.2	401	2 B90120	SNF1-related prote
29	561.5	41.8	513	1 S60303	serine/threonine-s

30	560.5	41.8	513	1 S60304	serine/threonine-s
31	541	40.3	891	2 T40503	protein kinase kin
32	540	40.2	891	2 A38903	protein kinase 1 -
33	534.5	39.8	441	2 C84667	probable protein k
34	533	39.7	651	2 S52244	p69Eg3 protein - A
35	532.5	39.7	480	2 A86427	probable serine/th
36	527.5	39.3	440	2 T14735	probable serine/th
37	526	39.2	1142	2 S59359	GIN4 protein - yea
38	525.5	39.2	440	2 T14736	probable serine/th
39	525	39.1	1147	2 S64930	serine/threonine-s
40	523	39.0	1064	2 S52687	serine/threonine-s
41	521	38.8	435	2 E84707	probable protein k
42	516	38.5	1518	2 S37928	probable purine nu
43	510	38.0	502	2 T02306	probable protein k
44	507.5	37.8	445	2 T09903	serine/threonine-s
45	500.5	37.3	421	2 E96522	hypothetical prote

ALIGNMENTS

RESULT 1

T29253  
hypothetical protein B0496.3 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C/Accession: T29253  
R/Murray, J.; Le, T.T.  
submitted to the EMBL Data Library, May 1996  
A/Description: The sequence of C. elegans cosmid B0496.  
A/Reference number: Z20596  
A/Accession: T29253  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1558 <MUR>  
A/Cross-references: EMBL:U58749; PIDN:AAB00636.1; GSPDB:GN00022; CESP:B0496.3  
A/Experimental source: strain Bristol N2; clone B0496  
C/Genetics:  
A/Gene: CESP:B0496.3  
A/Map position: 4  
A/Introns: 55/1; 100/3; 141/1; 226/3; 298/1; 468/3; 502/2; 609/3; 656/3; 795/3; 843/2; 91

Query Match	60.6%	Score	813.5	DB	2	Length	1558		
Best Local Similarity	61.7%	Pred. No.	2.7e-27						
Matches	156	Conservative	33	Mismatches	63	Indels	1	Gaps	1
Qy	1	YEFLETGKTYGKVKARESS-GRVVAIKSIRKDKIKDQDLHIREIEIEMSLNHPH	59						
Db	75	FEITKLGSGTYGKVSLAYDHFDEAVAVKLIKSAIESKADLVIRIRREIRIMSALNHPN	134						
Qy	60	IIAIHEVFENSSKIVIVMEYASRGDLYDIYSERPLSERDARHFFROIVSALHYCHONGI	119						
Db	135	IIQIYEFENKDKIILVMEYSSGGELYVSRGSLPEAGARRIFROITSVAVLYCHKHRV	194						
Qy	120	VHRDLKLENTLLDANGNIKIDFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYGVPEV	179						
Db	195	AHRDLKLENTLLDQNNAKIADFGLSNYFADKLLITFCGSPLYASPEIINGTYKQPEV	254						
Qy	180	DSWSGLVLLIYLVHGTWMPFDGQDHKTIVKOISNGAVREPPKPSDACLIRLLMVNPTTR	239						
Db	255	DCWSLGLTLLTVLYGSMFPDGRDNRMVRQIKRGAYFEPETPTSTASMLIRNMLRVNPERR	314						
Qy	240	ATLEDVASHWVN	252						
Db	315	ATIFDIASHWLN	327						

RESULT 2

G01025  
serine/threonine protein kinase - human  
C/Species: Homo sapiens (man)  
C/Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 16-Aug-2004  
C/Accession: G01025

R;Navarro, E.  
submitted to the EMBL Data Library, April 1996  
A:Reference number: H00564  
A:Accession: G01025  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-745 <NAV>  
A:Cross-references: UNIPROT:Q15524; UNIPROT:Q96HB3; EMBL:X97630; NID:g1310674  
C:Superfamily: protein kinase homology  
F:18-271/Domain: protein kinase homology <KIN>

Query Match 48.2%; Score 647.5; DB 2; Length 745;  
Best Local Similarity 51.2%; Pred. No. 1.3e-20;  
Matches 130; Conservative 44; Mismatches 77; Indels 3; Gaps 3;

Qy 1 YFELETGKGTGYGVKKARE-SSGRLVAIKRDKIKDEQDLHIREIEIMSSLNHPH 59  
Db 20 YRLKTKTGKGNFAKVLARHILTCGEVAVKIIDKTQL-NSSSLQKLFREVRIMKVLNHPN 78  
Qy 60 IIAIHEVFENSKKIVIMVEYASRGDLYISERPRLSERDARHFPRQIVSALHYCHONGI 119  
Db 79 IVKLFEVIEVTEKTLVLYMEYASGGEVFDYLVAGRMKEKARFQIVSAVQYCHOKFI 138  
Qy 120 VHRDLKLENTLLDANGNIKIADFGSLNLYHKGFLOTFCGSPLYASPEIVNGKYVGPVEV 179  
Db 139 VHRDLKAENLLDADNMKIADFGSNEFTFGNKLDITFCGSPPYAAPLFGKKYDGPVEV 198  
Qy 180 DWSLGLVLLYLVHGTMPFGQDHKTLVKQISNGAYREPPKPSDAC-GLIRWLLMNPTR 238  
Db 199 DVWSLGLVILYLVSGSLPFDGQNLKELRERVLGRKYRIPFYMSTDCENLLKFLILNPSK 258  
Qy 239 RATLEDVASHWVN 252  
Db 259 RGTLEQIMKDRWIN 272

RESULT 3  
S27966  
probable serine/threonine-specific protein kinase (EC 2.7.1.-) - human  
N:Alternate names: protein p78  
C:Species: Homo sapiens (man)  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 16-Aug-2004  
A:Accession: S27966  
R;Maheshwari, K.K.; Som, S.; Parsa, I.  
submitted to the EMBL Data Library, January 1992  
A:Description: Sequence of a cDNA encoding 78kD marker protein lost in chemically induced  
A:Reference number: S27966  
A:Molecule type: mRNA  
A:Residues: 1-713 <MAH>  
A:Cross-references: UNIPROT:P27448; EMBL:M80359; NID:g189511; PIDN:AAA59991.1; PID:g1895  
C:Superfamily: protein kinase homology  
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
F:54-307/Domain: protein kinase homology <KIN>  
F:62-70/Region: protein kinase ATP-binding motif

Query Match 48.2%; Score 646.5; DB 2; Length 713;  
Best Local Similarity 51.6%; Pred. No. 1.4e-20;  
Matches 131; Conservative 41; Mismatches 79; Indels 3; Gaps 3;

Qy 1 YFELETGKGTGYGVKKARE-SSGRLVAIKRDKIKDEQDLHIREIEIMSSLNHPH 59  
Db 56 YRLKTKTGKGNFAKVLARHILTCGEVAVKIIDKTQL-NPTSLQKLFREVRIMKVLNHPN 114  
Qy 60 IIAIHEVFENSKKIVIMVEYASRGDLYISERPRLSERDARHFPRQIVSALHYCHONGI 119  
Db 115 IVKLFEVIEVTEKTLVLYMEYASGGEVFDYLVAGRMKEKARFQIVSAVQYCHOKRI 174  
Qy 120 VHRDLKLENTLLDANGNIKIADFGSLNLYHKGFLOTFCGSPLYASPEIVNGKYVGPVEV 179  
Db 175 VHRDLKAENLLDADNMKIADFGSNEFTVGKLDITFCGSPPYAAPLFGKKYDGPVEV 234  
Qy 180 DWSLGLVLLYLVHGTMPFGQDHKTLVKQISNGAYREPPKPSDAC-GLIRWLLMNPTR 238

Db 235 DVWSLGLVILYLVSGSLPFDGQNLKELRERVLGRKYRIPFYMSTDCENLLKFLILNPIK 294  
Qy 239 RATLEDVASHWVN 252  
Db 295 RGTLEQIMKDRWIN 308

RESULT 4  
T52633  
serine/threonine-specific protein kinase (EC 2.7.1.-) AKIN11 [validated] - Arabidopsis th  
N:Alternate names: SNF1 protein kinase omolog AKIN11  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 09-Jul-2004  
A:Accession: T52633  
R;Bhalerao, R.P.; Salchert, K.; Bako, L.; Okresz, L.; Szabados, L.; Muranaka, T.; Machida  
PROG. Natl. Acad. Sci. U.S.A. 96, 5322-7, 1999  
A>Title: Regulatory interaction of PRL1 WD protein with Arabidopsis SNF1-like protein kin  
A:Reference number: Z25116; MUID:99238528; PMID:10220464  
A:Accession: T52633  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-512 <BHA>  
A:Cross-references: UNIPROT:P92958; EMBL:X99279; PIDN:CAA67671.1  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Gene: AKIN11  
C:Function:  
A:Description: EC 2.7.1.-; serine/threonine-specific protein kinase AKIN11 [validated, M  
complements SNF1 mutations in yeast  
C:Superfamily: AMP-activated protein kinase; protein kinase homology  
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

Query Match 48.0%; Score 644; DB 2; Length 512;  
Best Local Similarity 50.8%; Pred. No. 1.4e-20;  
Matches 128; Conservative 47; Mismatches 75; Indels 2; Gaps 2;

Qy 1 YFELETGKGTGYGVKKARE-SSGRLVAIKRDKIKDEQDLHIREIEIMSSLNHPH 59  
Db 20 YLKGKTLGIGSGFVKIAEHVVTGHVAKILNRRKIKNMEEKVREIKILRFLMHPH 79  
Qy 60 IIAIHEVFENSKKIVIMVEYASRGDLYISERPRLSERDARHFPRQIVSALHYCHONGI 119  
Db 80 IIRQVEVETTSDIYVMEYVKGSELFYIVEKGRLEQDEARNFFQIIISGVEYCHENMV 139  
Qy 120 VHRDLKLENTLLDANGNIKIADFGSLNLYHKGFLOTFCGSPLYASPEIVNGKYVGPVEV 179  
Db 140 VHRDLKPENLLDSRCNIKIADFGLSNVNRDGHFLKTSCTGSPNYAAPEVISGKLYAGPEV 199  
Qy 180 DWSLGLVLLYLVHGTMPFGQDHKTLVKQISNGAYREPPK-PSDACGLIRWLLMNPTR 238  
Db 200 DVWSGCVILYALCGILPFDENIPNLFKKIKGGIYTLFSLHSSEARDLIPRLIIVDPVK 259  
Qy 239 RATLEDVASHWVN 250  
Db 260 RITPIRQHRW 271

RESULT 5  
T10449  
probable serine/threonine-specific protein kinase (EC 2.7.1.-) - cucumber  
N:Alternate names: SNF1-related protein kinase  
C:Species: Cucumis sativus (cucumber)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
A:Accession: T10449  
R;Gumpel, N.J.  
submitted to the EMBL Data Library, December 1996  
A:Reference number: Z17020  
A:Accession: T10449  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-504 <GUM>  
A:Cross-references: UNIPROT:P93113; EMBL:Y10036

Db 259 RVTIPEIRQHPW 270

Qy 239 RATLEDVASHWTV 251

RESULT 8  
A56009  
serine/threonine-specific protein kinase (EC 2.7.1.1-) NPK5 - common tobacco  
C;Species: Nicotiana tabacum (common tobacco)  
C;Date: 28-Apr-1995 #sequence\_revision 28-Apr-1995 #text\_change 09-Jul-2004  
C;Accession: A56009  
R;Muranaka, T.; Banno, H.; Machida, Y.  
Mol. Cell. Biol. 14, 2958-2965, 1994  
A;Title: Characterization of tobacco protein kinase NPK5, a homolog of Saccharomyces cerevisiae  
A;Reference number: A56009; MUID:94217693; PMID:8164654  
A;Accession: A56009  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-511 <MUR>  
A;Cross-references: UNIPROT:Q40544; GB:D26602; NID:G496384; PIDN:BAA05649.1; PID:G496385  
C;Function:  
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin  
C;Superfamily: AMP-activated protein kinase; protein kinase homology  
C;Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase  
F;11-271/Domain: protein kinase homology <KIN>  
F;25-33/Region: protein kinase ATP-binding motif  
F;48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted  
F;147,151/Binding site: magnesium (Asn, Asp) #status predicted  
Query Match 47.0%; Score 631; DB 1; Length 511;  
Best Local Similarity 49.6%; Pred. No. 4.7e-20;  
Matches 125; Conservative 48; Mismatches 77; Indels 2; Gaps 2;  
Qy 1 YEFLETGKGYGVKKARE-SGRVVAIKSIRKDKIDQDLHIREIEIMSSLNHPH 59  
Db 19 YLKGKTIGISFGVKVTAETHLTGHVAVKILNRKIKNMEMEKVREIKILFPHPH 78  
Qy 60 IIAIHEVFENSSKIVIMVEYASRGDLVDYISERPRLSERDARHPFRQIVSALHYCHQNGI 119  
Db 79 IIRLYEVETPSDIYVMEYVYKSGELFDYIEVKRLQDEARKFFQIISGVEYCHRMV 138  
Qy 120 VHRDLKLENTLLDANGNIKIADFGLSNLYHKGFLOFTFCGSPLYASPEIVNGPKYGVPEV 179  
Db 139 VHRDLKPENLDDSKNVKIKIADFGLSNIMRDGHFLKTS CGSPNVAAPAEVISGKIYAGPEV 198  
Qy 180 DWSLGLVLLYLVTGTMPPFDGQDHKTLVKQISNGAYREPPKPS-DACGLIRWLLMVNPT 238  
Db 199 DVWSGVLVALLCGTLPFDDEINPFLFKIKGGWISPSHLSAGARDLIPRLMILVDPMK 258  
Qy 239 RATLEDVASHMW 250  
Db 259 RWTPEIRMPHW 270  
RESULT 9  
I48609  
probable serine/threonine-specific protein kinase (EC 2.7.1.1-) kem - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 16-Aug-2004  
C;Accession: I48609; S31333  
R;Ingilis, J.D.; Lee, M.; Hill, R.E.  
Mamm. Genome 4, 401-403, 1993  
A;Title: Emk, a protein kinase with homologs in yeast maps to mouse chromosome 19.  
A;Reference number: I48609; MUID:93364122; PMID:8358177  
A;Accession: I48609  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-774 <INGI>  
A;Cross-references: UNIPROT:Q05512; EMBL:X70764; NID:G57919; PIDN:CAA50040.1; PID:G57920  
R;Ingilis, J.D.; Lee, M.; Hill, R.E.  
submitted to the EMBL Data Library, January 1993  
A;Description: A novel protein kinase with homologues in yeast maps to mouse chromosome  
A;Reference number: S31333  
A;Accession: S31333  
A;Molecule type: mRNA  
A;Residues: 1-698, 'K', 700-702, 'GRGLPTAAKASCTSGTCYCAGMAHQATRT', 731-774 <ING2>

A;Cross-references: EMBL:X70764

C;Genetics:

A;Gene: emk

C;Superfamily: protein kinase homology

C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F;51-304/Domain: protein kinase homology &lt;KIN&gt;

F;59-67/Region: protein kinase ATP-binding motif

Query Match 46.8%; Score 627.5; DB 2; Length 774;

Best Local Similarity 50.0%; Pred. No. 8.9e-20;

Matches 127; Conservative 44; Mismatches 80; Indels 3; Gaps 3;

Qy 1 YEFLETGKGYGVKKARE-SGRVVAIKSIRKDKIDQDLHIREIEIMSSLNHPH 59

Db 53 YRLKLTIGKNFAKVKLRHILTGVEVAVKIITQTL-NSSSLQKLPREVIMKVLNHPN 111

Qy 60 IIAIHEVFENSSKIVIMVEYASRGDLVDYISERPRLSERDARHPFRQIVSALHYCHQNGI 119

Db 112 IVKLFEVIEETKTLVLMVEYASGGEVFDYLVAHGRMKEKARAKFRQIVLHVQYCHQKFI 171

Qy 120 VHRDLKLENTLLDANGNIKIADFGLSNLYHKGFLOFTFCGSPLYASPEIVNGPKYGVPEV 179

Db 172 VHRDLKAENLLDADNMNIKIADFGFSNEFTFGNKLDTFCGSPPYAAPDELFGKKIDGPEV 231

Qy 180 DWSLGLVLLYLVTGTMPPFDGQDHKTLVKQISNGAYREPPKPSDAC-GLIRWLLMVNPT 238

Db 232 DVWSLGVILYTLVSGSLFFDQGNLKELRERVLGRKRIPIYMTDCENLLKLFILNPSK 291

Qy 239 RATLEDVASHMWVN 252

Db 292 RGTLEQIMKDRWN 305

## RESULT 10

T18611

probable serine/threonine-specific protein kinase (EC 2.7.1.1-), long splice form - Caenorhabditis elegans

N;Contains: probable serine/threonine kinase, short splice form

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T18610; T23144; T23143

R;McMurray, A.

submitted to the EMBL Data Library, October 1996

A;Reference number: Z18997

A;Accession: T18611

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1192 &lt;WILI&gt;

A;Cross-references: UNIPROT:Q9TW45; EMBL:Z81027; PIDN:CAB54179.1; GSPDB:GN00023; CESP:H39E23.1b

A;Experimental source: clone AH10

A;Accession: T18610

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-487, 536-1192 &lt;WIL2&gt;

A;Cross-references: EMBL:Z81027; PIDN:CAB54178.1; GSPDB:GN00023; CESP:H39E23.1b

A;Experimental source: clone AH10

R;McMurray, A.

submitted to the EMBL Data Library, June 1997

A;Reference number: Z19696

A;Accession: T23144

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1192 &lt;WIL3&gt;

A;Cross-references: EMBL:Z96102; PIDN:CAB54263.1; GSPDB:GN00023; CESP:H39E23.1a

A;Experimental source: clone H39E23

A;Accession: T23143

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-487, 536-1192 &lt;WIL4&gt;

A;Cross-references: EMBL:Z96102; PIDN:CAB54262.1; GSPDB:GN00023; CESP:H39E23.1b

A;Experimental source: clone H39E23

C;Genetics:

A;Gene: CESP:H39E23.1a; CESP:H39E23.1b

A;Map position: 5



A; Introns: 10/3; 33/3; 67/3; 139/2; 183/3; 451/3; 487/3; 535/3; 631/1; 825/2; 914/3; 952/3  
C; Keywords: alternative splicing; ATP; phosphotransferase; serine/threonine-specific protein kinase  
F; 1-1192/Product: probable serine/threonine-specific protein kinase, long splice form #1  
F; 1-487,536-1192/Product: probable serine/threonine-specific protein kinase, short splice form #2

Query Match 46.3%; Score 621.5; DB 2; Length 1192;  
Best Local Similarity 48.0%; Pred. No. 2.2e-19;  
Matches 122; Conservative 52; Mismatches 77; Indels 3; Gaps 3;

Qy 1 YEFLETIGKTYGVKKVKAARE-SGRLVAIKSIRKDIIKDQDLHIRREIEIMSSLNHPH 59  
Db 170 YKLKTIKGNGFAKXVIKAKHVITCHEVAIKIIDKTAL-NPSSLQKLFREVKIMKQLDHPN 228  
Qy 60 IIAIHEVPENSKIVIMEVASRGDLVDYISERPLSERDARHFFROIVSAHYCHONGI 119  
Db 229 IVKLYQWETEQTLYLVEIASGEVFDYLVAHGMEKEARAKFRQIVSAVOYLHSKNI 288  
Qy 120 VHRDLKLEILLDANGNIKIADFGSLNLVHKGFLOTFCGSPLYASPEIVNGKPYPVGPEV 179  
Db 289 IHRDLKAENLLDDMMIKIADFSGFTNSLGNNKLTDFCGSPPYAAPELPSGGKYDGPEV 348  
Qy 180 DWSGLGVLLYILVHGTWPFDQQDHKTLVKQISNGAYREP PPKPSDAC-GLIRWLMLVNPT R 238  
Db 349 DVMSGLVILYLVSGSLPFDFQNLRERVLGRKYRIPFYMTSDCNLLKKFLVINPQR 408  
Qy 239 RATLEDVASHWNVN 252  
Db 409 RSSLDNMKDRWMN 422

RESULT 11  
G89287  
C; Species: Caenorhabditis elegans  
C; Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C; Accession: G89287  
R; anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A; Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes  
A; Reference number: A75000; MUID:9089613; PMID:9851916  
A; Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/c\_elegans/  
A; Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999.  
A; Accession: G89287  
A; Status: preliminary  
A; Molecule type: DNA  
A; Residues: 1-1246 <STO>  
A; Cross-references: GB:Chr\_V; PIDN:CAB09532.1; PID:Jg3878100; GSPDB:GN00023; CESP:H39E23.1  
C; Genetics:  
A; Gene: H39E23.1  
A; Map position: 5

Query Match	46.3%	Score	621.5;	DB 2;	Length	1246;			
Best Local Similarity	48.0%;	Pred. NO.	2.2e-19;						
Matches	122;	Conservative	52;	Mismatches	77;	Indels	3;	Gaps	3;
Qy	1	YEFLETGKGYGVKKKARE--SSGRLVAIKSIRKDKIKDEQDLLHIREIEMSLNHPH	59						
Db	60	YKLKLTGKGNFAVKLAKHVITGHEVAIKLIDTKAL--NPSSLQKLFREVMKQLQDHPN	118						
Qy	60	IIATHEFYFENSKKIVIMEYASRGDLVDYISERPLSRDARHPFRQIVSAHLHYCHQNGI	119						
Db	119	IVKLYQWMEQTQLVLYLEVYASGEVFPDYLVAHGRMKEKARAFQIVSAQVYLHSKNI	178						
Qy	120	VHRDLKLENILLDANGHIKIIDAFGLSNLYHKGKFLQTPCGSPLYASPIVNGKPYVGDEV	179						
Db	179	IHRDLKAENLLLDQMMIKIIDAFGFSNTFSLGNKLDTCFSGPPYAAPELFSGKYDGPV	238						
Qy	180	DSWSLGYLLYLTVHGTMPFDGQDHKTIVKQJLSNGAYREPPKPSDAC--GLISLLLMVNPT	238						
Db	239	DVWSLGVILYTLVSGSPFPDQGNLKELRERVLRGKYRIPPFYMTDCENLLKFLVINPQR	298						
Qy	239	RATLEDVASHWVN	252						

Df  
299 RSSLDNIMKDRWN 312

RESULT 12  
B90100  
SNF-related kinase [imported] - Guillardia theta nucleomorph  
C:Species: nucleomorph Guillardia theta  
A>Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: B90100  
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reif  
Nature 410, 1091-1096, 2001  
A>Title: The highly reduced genome of an enslaved algal nucleus.  
A:Reference number: A39082; PMID:11323671; PMID:11323671  
A:Accession: B90100  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-472 <DOU>  
A:Cross-references: UNIPROT:Q98RL9; GB:AF165818; NID:g13794554; PIDN:AAK39929.1; GSPDB:G:  
C:Genetics:  
A:Gene: kin(enf1)  
A:Map position: 1  
A:Genome: nucleomorph  
C:Superfamily: AMP-activated protein kinase; protein kinase homology  
C:Keywords: nucleomorph

Query Match 46.3%; Score 621; DB 2; Length 472;  
Best Local Similarity 48.0%; Pred. No. 1.2e-19;  
Matches 122; Conservative 52; Mismatches 78; Indels 2; Gaps 2;

Qy 1 YEFLETLGKGTYGVKKKA-RESSGRLLVAIKSRKDQDEQLLHRRREIEIMSSLNHPH 59  
| : || | : || | : || | : || | : || | : || | : || | : || | : ||  
Db 12 YYLGKTGLVGSFGVKLGEHELCGKVAVKLNRKKIKNLMKEEKVKREICILKLFMHPH 71  
| : || | : || | : || | : || | : || | : || | : || | : || | : ||  
Qy 60 IIAIHVEPENSKIIVMEYASRGDLYDISRPISRDLSEARHFHQIVSALHYCHONGI 119  
| : || | : || | : || | : || | : || | : || | : || | : || | : ||  
Db 72 IIRLEYEVETPDIFPVVTYEITGGELFDYIVERGRNLNEDESRRKFQQMGISIEYCHNHV 131  
| : || | : || | : || | : || | : || | : || | : || | : || | : ||  
Qy 120 VHRDLKENILLDANGNTIKADFGLSNLYHKGFQTFCGSLPYASPEIVNGKPVGPEV 179  
| : || | : || | : || | : || | : || | : || | : || | : || | : ||  
Db 132 VHRDLKPENILLDAHLNVKIADFGLSNIMKOGNFLTKSCGPNYAAPVINGKSYLGPV 191  
| : || | : || | : || | : || | : || | : || | : || | : || | : ||  
Qy 180 DWSLGLVLLYYLVHTGMTFPDGDDHKTVLKQISNGAYREPPEPSD-ACGLRWLLMVNPTR 238  
| : || | : || | : || | : || | : || | : || | : || | : || | : ||  
Db 192 DVWCSGVIMYALLCGSLPFDDDENINPLFKKIGIYILPGYLSDLSDRMIAKMLTTPLL 251  
| : || | : || | : || | : || | : || | : || | : || | : || | : ||  
Qy 239 RATLEDVASHWVN 252  
| : || | : || | : || | : || | : || | : || | : || | : || | : ||  
Db 252 RITINEIRDHPWFN 265  
| : || | : || | : || | : || | : || | : || | : || | : || | : ||

RESULT 13  
A26030  
serine/threonine-specific protein kinase [EC 2.7.1.-) SNF1 - yeast (Saccharomyces cerevisiae)]  
N:Alternate names: protein YDR477w  
C:Species: Saccharomyces cerevisiae  
C>Date: 20-Aug-1987 #sequence\_revision 20-Aug-1987 #text\_change 09-Jul-2004  
R:Ceelenza, J.L.; Carlson, M.  
Science 233, 1175-1180, 1986  
A>Title: A yeast gene that is essential for release from glucose repression encodes a pr  
A:Reference number: A26030; PMID:86289463; PMID:3526554  
A:Accession: A26030  
A:Molecule type: DNA  
A:Residues: 1-633 <CEL>  
A:Cross-references: UNIPROT:P06782; EMBL:M13971; NID:g172629; PIDN:AAA35058.1; PID:g1726  
R:Dieckrich, F.S.  
submitted to the EMBL Data Library, August 1995  
A>Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.  
A:Reference number: S69554  
A:Accession: S69644  
A:Molecule type: DNA  
A:Residues: 1-633 <DIE>

A;Cross-references: EMBL:U33050; NID:g927726; PIDN:AAB64904.1; PID:g927732; GSPDB:GN0000  
C;Genetics:  
A;Gene: SGD:SNF1; MIPS:YDR477w  
A;Cross-references: SGD:S0002885; MIPS:YDR477w  
A;Map position: 4R  
C;Function:  
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin  
A;Note: required for expression of glucose-repressed genes in response to glucose depriv  
C;Superfamily: AMP-activated protein kinase; protein kinase homology  
C;Keywords: ATP; autophosphorylation; magnesium; nucleus; phosphoprotein; phosphotransfe  
F;53-306/Domain: protein kinase homology <KIN>  
F;61-69/Region: protein kinase ATP-binding motif  
F;84,103,177,179/Active site: Lys, Glu, Asp, Lys #status predicted  
F;182,186/Binding site: magnesium (Asn, Asp) #status predicted  
F;210/Binding site: phosphate (Thr) (covalent) (by unidentified kinase) #status predicted

Query Match 46.2%; Score 619.5; DB 1; Length 633;  
Best Local Similarity 48.2%; Pred. No. 1.7e-19;  
Matches 122; Conservative 57; Mismatches 69; Indels 5; Gaps 5;

Qy 1 YFELETGKTYGKVKKA-RESSGLVAIKSKDKIKDEODLL-HIRREIEMSSLNHP 58  
Db 55 YQIVKTLGEGFGKYLAYHTTTGQVALKIINK-KVLAKSDMQGRIEREISYLLRLRHP 113

Qy 59 HIIAHEVFENSKIVIVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONG 118  
Db 114 HIIKLDYIVKSKDELIWIEVAG-NELPDYIVQDKKSEQARRFPQIISAVEYCHRHK 172

Qy 119 IVHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYVGPE 178  
Db 173 IVHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYVGPE 232

Qy 179 VDSWSGLVLLYLHVGTMPPFGQDHKTLVKQISNGAYREPPEKPS-DACGLIRWLLMNPTR 237  
Db 233 VDVMSGCVILYVLCRLRPFDDSEIPVLFKNISNGVYTLPKFLSPGAAGLIKRLMILVNP 292

Qy 238 RRTLEDVASHHW 250  
Db 293 NRISIHEIMQDDW 305

RESULT 14  
T07415  
probable serine/threonine-specific protein kinase (EC 2.7.1.-) PKIN1 - potato  
N;Alternate names: SNF1-related protein kinase  
C;Species: Solanum tuberosum (potato)  
C;Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004  
C;Accession: T07415  
R;Halford, N.G.  
submitted to the EMBL Data Library, March 1996  
A;Reference number: Z16022  
A;Accession: T07415  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-504 <HAL>  
A;Cross-references: UNIPROT:Q41485; EMBL:X95997; NID:g1216279; PIDN:CAA65244.1  
C;Genetics:  
A;Gene: PKIN1  
A;Introns: 62/1; 123/3; 184/3; 228/3; 290/3; 320/3; 348/3; 389/3; 466/3  
C;Function:  
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin  
C;Superfamily: AMP-activated protein kinase; protein kinase homology  
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
F;15-269/Domain: protein kinase homology <KIN>

Query Match 46.1%; Score 618; DB 2; Length 504;  
Best Local Similarity 48.8%; Pred. No. 1.6e-19;  
Matches 123; Conservative 45; Mismatches 82; Indels 2; Gaps 2;

Qy 1 YFELETGKTYGKVKKARE-SSGRIVATKSRKDKIKDEODLLHIRREIEMSSLNHP 59  
Db 17 YRVGKTLGHSFGKVAIEHLHTGHKVAIKILRRRKMTDPMEELRREIKRLFLVPH 76

Qy 60 IIAIHEVFENSKIVIVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONG 119  
Db 77 VLRLEVIETPTDIYVMEYVKSGELFDYIVKGRLOEAEARKIFQIIAGVEYCHKMV 136

Qy 120 VHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYVGPE 179  
Db 137 VHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYVGPE 196

Qy 180 DWSLSGLVLLYLHVGTMPPFGQDHKTLVKQISNGAYREPPEKPSD-ACGLIRWLLMNPTR 238  
Db 197 DVMSGCVILYVLCRLRPFDDSEIPVLFKNISNGVYTLPKFLSPGAAGLIKRLMILVDP 256

Qy 239 RATLEDVASHHW 250  
Db 257 RISVPDIRQHOW 268

RESULT 15  
JC7500  
qik protein - chicken  
N;Alternate names: Qin-induced kinase  
C;Species: Gallus gallus (chicken)  
C;Date: 17-Nov-2000 #sequence\_revision 17-Nov-2000 #text\_change 09-Jul-2004  
C;Accession: JC7500  
R;Xia, Y.; Zhang, Z.; Kruse, U.; Vogt, P.K.; Li, J.  
Biochem. Biophys. Res. Commun. 276, 564-570, 2000  
A;Title: The new serine-threonine kinase, Qik, is a target of the qin oncogene.  
A;Reference number: JC7500  
A;Contents: Embryo fibroblasts  
A;Accession: JC7500  
A;Molecule type: mRNA  
A;Residues: 1-798 <XIA>  
A;Cross-references: UNIPROT:Q91A88; GB:AF219232  
C;Comment: This protein, a member of the AMPK/SNF1 family of serine/threonine kinases, f  
C;Genetics:  
A;Gene: qik  
C;Keywords: protein kinase

Query Match 46.0%; Score 617.5; DB 2; Length 798;  
Best Local Similarity 48.6%; Pred. No. 2.4e-19;  
Matches 123; Conservative 47; Mismatches 80; Indels 3; Gaps 3;

Qy 1 YFELETGKTYGKVKKARESSGRL-VAIKSRKDKIKDEODLLHIRREIEMSSLNHP 59  
Db 26 YDIERTLGKGNFAVVKLARHRTKTQVAKIIDKTRL-DPSNLEKIYREVOIMKLLNHP 84

Qy 60 IIAIHEVFENSKIVIVMEYASRGDLVDYISERPRLSERDARHFRQIVSALHYCHONG 119  
Db 85 IIKLYQWETKMDLIYVTEFAKNGEMFDHLTSNGHLSSEARKKFWQLISAVEYCHSHI 144

Qy 120 VHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYVGPE 179  
Db 145 VHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIYNGKPYVGPE 204

Qy 180 DWSLSGLVLLYLHVGTMPPFGQDHKTLVKQISNGAYREPPEKPSDAC-GLIRWLLMNPTR 238  
Db 205 DWSLSGLVLLYLHVGTMPPFGQDHKTLVKQISNGAYREPPEKPSDAC-GLIRWLLMNPTR 264

Qy 239 RATLEDVASHHW 251  
Db 265 RITISQIKQKHW 277

Search completed: January 22, 2005, 06:43:52  
Job time : 16.4548 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 06:28:10 ; Search time 74.4118 Seconds  
(without alignments)  
1956.274 Million cell updates/sec

Title: US-09-980-464-11\_COPY\_57\_309

Perfect score: 1342

Sequence: 1 YEFLETLGKTYGKVKARE.....VNPRTRATLEDVASHVWVNW 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1342	100.0	631	2 Q8C1C0	Q8c1c0 mus muscu
2	1342	100.0	631	2 Q9DBV0	Q9dbv0 m mus muscu
3	1328	99.0	639	2 Q80ZW3	Q80zw3 mus muscu
4	1328	99.0	639	2 Q8BZNA	Q8bzn4 mus muscu
5	1293	96.3	628	2 Q9H093	Q9h093 homo sapien
6	1128.5	84.1	661	1 AAK5 HUMAN	O60285 homo sapien
7	1128.5	84.1	698	2 BAA25463	Baa25463 homo sapi
8	1025	76.4	575	2 Q616D6	Q616d6 mus muscu
9	824.5	61.4	383	2 Q7QBH4	Q7qbh4 anopheles g
10	822.5	61.3	1180	2 Q7KSS0	O7kss0 drosophila
11	822.5	61.3	1180	2 Q7YUL9	O7yul9 drosophila
12	822.5	61.3	1180	2 AAS65135	Aas65135 drosophil
13	822.5	61.3	1427	2 Q9VH05	Q9vh05 drosophila
14	822.5	61.3	1427	2 AAF54517	Aaf54517 drosophil
15	813.5	60.6	1566	2 Q8WQG7	Q8wqg7 caenorhabdi
16	813.5	60.6	1607	2 Q9B1B1	Q9b1b1 caenorhabdi
17	665.5	49.6	752	1 MKK4 HUMAN	Q96134 homo sapien
18	665.5	49.6	752	2 Q8C1P4	Q8c1p4 mus muscu
19	660.5	49.2	752	2 Q8NG37	Q8ng37 homo sapien
20	652.5	48.6	725	2 Q804T2	Q804t2 xenopus lae
21	651.5	48.5	729	2 Q9JKE4	Q9jke4 mus muscu
22	651.5	48.5	744	2 Q9JKE5	Q9jke5 mus muscu
23	650	48.4	712	2 Q7QBH3	Q7qbh3 anopheles g
24	647.5	48.2	691	2 Q96RG0	Q96rg0 homo sapien
25	647.5	48.2	722	2 Q08679	O08679 rattus norv
26	647.5	48.2	722	2 Q08679	O08679 rattus norv
27	647.5	48.2	722	2 AAH58556	Aah58556 mus muscu
28	647.5	48.2	725	2 Q15524	Q15524 homo sapien
29	647.5	48.2	755	2 Q7KZ17	Q7kz17 homo sapien
30	647.5	48.2	778	2 Q96BH3	Q96bh3 homo sapien
31	647.5	48.2	778	2 AAH08771	Aah08771 homo sapi

32 647.5 48.2 797 2 Q8VHF0 Q8vhf0 rattus norv  
33 647.5 48.2 888 2 Q8BR95 Q8br95 mus muscu  
34 644 48.0 512 2 P92958 P92958 arabidopsis  
35 644 48.0 512 2 P92958 P92958 arabidopsis  
36 643.5 48.0 722 2 Q802W0 Q802w0 brachydanio  
37 642.5 47.9 792 2 Q61NT7 Q6int7 xenopus lae  
38 642.5 47.9 793 2 AAH72186 Aah72186 xenopus l  
39 642.5 47.9 793 2 O08678 O08678 rattus norv  
40 642.5 47.9 795 2 Q8VHJ5 Q8vhj5 mus muscu  
41 641.5 47.8 776 2 Q7ZYL7 Q7zyl7 xenopus lae  
42 641.5 47.8 785 2 Q8QGV3 Q8qgv3 xenopus lae  
43 641 47.8 504 2 P93113 P93113 cucumis sat  
44 640.5 47.7 462 2 O75271 O75271 homo sapien  
45 640.5 47.7 542 2 Q6V8Y5 Q6v8y5 physcomitre

## ALIGNMENTS

RESULT 1  
Q8C1C0 ID Q8C1C0 PRELIMINARY; PRT; 631 AA.  
AC Q8C1C0;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE 1200013B22Rik protein.  
GN Name=1200013B22Rik;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Czech II;  
RC TISSUE=Mammary tumor metastatized to lung. MMTV-LTR/Wnt1 model.  
RC Expression driven by an MMTV-LTR enhancer.;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez A.,  
RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,  
RA Blakesley R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Czech II;  
RC TISSUE=Mammary tumor metastatized to lung. MMTV-LTR/Wnt1 model.  
RC Expression driven by an MMTV-LTR enhancer.;  
RA Strausberg R.,  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
DR EMBL; BC033302; AAH33302.1; -.  
DR HSSP; P31751; 1GZK.  
DR MGD; MGI:1921387; 1200013B22Rik  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

DR InterPro; IPR011009; Kinase like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 631 AA; 69745 MW; C6C9042247D9174F CRC64;

Query Match 100.0%; Score 1342; DB 2; Length 631;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-97;  
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YEFLETGKGYGVKVKARESSGRLVAIKSRDKIKDEQDLLHRIEIEIMSSLNPHI 60  
 Db 57 YEFLETGKGYGVKVKARESSGRLVAIKSRDKIKDEQDLLHRIEIEIMSSLNPHI 116  
 Qy 61 IAIHEVFENSKIIVMEYASRGDLYDISRPLSERDARHFRQIVSALHYCHQNGIV 120  
 Db 117 IAIHEVFENSKIIVMEYASRGDLYDISRPLSERDARHFRQIVSALHYCHQNGIV 176  
 Qy 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPELVNGKPYVGVPEVD 180  
 Db 177 HRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPELVNGKPYVGVPEVD 236  
 Qy 181 SWSLGVLLYLLVHGTMPDGDQDHKTIVKQISNGAYRPPKPSDAGLIRLLMVNPTTRA 240  
 Db 237 SWSLGVLLYLLVHGTMPDGDQDHKTIVKQISNGAYRPPKPSDAGLIRLLMVNPTTRA 296  
 Qy 241 TLEDVASHWVNW 253  
 Db 297 TLEDVASHWVNW 309

RESULT 2  
 Q9DBV0 PRELIMINARY; PRT; 631 AA.  
 ID Q9DBV0;  
 AC Q9DBV0;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Mus musculus adult male lung cDNA, RIKEN full-length enriched library,  
 DE clone:1200013B22 product:weakly similar to PROBABLE SERINE/THREONINE-  
 DE PROTEIN KINASE KIAA0537 (EC 2.7.1.-) (Mus musculus adult male cecum  
 DE cDNA, RIKEN full-length enriched library, clone:9130215K18  
 DE product:weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE  
 DE KIAA0537)  
 GN Name=1200013B22Rik;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;

RA The FANTOM Consortium.  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Cecum, and Lung;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Taehiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Lung;  
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai K., Sakai K.,  
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Tejima Y., Toya T., Yamamoto T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Cecum;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kaga I., Kasukawa T.,  
 RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 DR EMBL; AK004737; BAB23518.1; -  
 DR EMBL; AK033672; BAC28421.1; -  
 DR HSSP; P31751; 1GZK.  
 DR MGD; MGI:1921387; 1200013B22Rik.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0005524; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR011009; Kinase-like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 1.

DR SMART; SM00220; S\_TKC; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 631 AA; 69806 MW; D51C042DE6CC174B CRC64;

Query Match 100.0%; Score 1342; DB 2; Length 631;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-97;  
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEFLETGKGTGKGVKKARESSGRLVAIKSIRKDKINDEQDLHIREIEMSSLNPHI 60  
 DB 57 YEFLETGKGTGKGVKKARESSGRLVAIKSIRKDKINDEQDLHIREIEMSSLNPHI 116  
 QY 61 IATHEVFENSKIVIVMEYASRGDLYISERPLSERDARHFFRQIVSALHCHONGIV 120  
 DB 117 IATHEVFENSKIVIVMEYASRGDLYISERPLSERDARHFFRQIVSALHCHONGIV 176  
 QY 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYVGPEVD 180  
 DB 177 HRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYVGPEVD 236  
 QY 181 SWSLGLVLLYLVHGTMPFDGQDKHTLVKQISNGAYRPPKPSDACGLIRLWLLMVNPTTRA 240  
 DB 237 SWSLGLVLLYLVHGTMPFDGQDKHTLVKQISNGAYRPPKPSDACGLIRLWLLMVNPTTRA 296  
 QY 241 TLEDVASHHWVNW 253  
 DB 297 TLEDVASHHWVNW 309

RESULT 3  
 Q802W3 PRELIMINARY; PRT; 639 AA.

AC Q802W3  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE 1200013B22Rik protein.  
 GN Name=1200013B22Rik;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skaleka U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX Strausberg R.;

Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 EMBL; BC046833; AAH46833.1; --  
 DR HSSP; P31751; ICZK.  
 DR MGD; MGI:1921387; 1200013B22Rik.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR011009; Kinase-like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR022290; Ser\_thr\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_kinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 639 AA; 70675 MW; FB9C40228F53872C CRC64;

Query Match 99.0%; Score 1328; DB 2; Length 639;  
 Best Local Similarity 96.9%; Pred. No. 1.6e-96;  
 Matches 253; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 YEFLETGKGTGKGVKKARESSGRLVAIKSIRKDKINDEQDLHIREIEMSSLNPHI 60  
 DB 57 YEFLETGKGTGKGVKKARESSGRLVAIKSIRKDKINDEQDLHIREIEMSSLNPHI 116  
 QY 61 IATHEVFENSKIVIVMEYASRGDLYISERPLSERDARHFFRQIVSALH 112  
 DB 117 IATHEVFENSKIVIVMEYASRGDLYISERPLSERDARHFFRQIVSALH 176  
 QY 113 YCHONGIVHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGK 172  
 DB 177 YCHONGIVHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGK 236  
 QY 173 PYVGPEVDSWSLGLVLLYLVHGTMPFDGQDKHTLVKQISNGAYRPPKPSDACGLIRWLL 232  
 DB 237 PYVGPEVDSWSLGLVLLYLVHGTMPFDGQDKHTLVKQISNGAYRPPKPSDACGLIRWLL 296  
 QY 233 MVNPTTRATLEDVASHHWVNW 253  
 DB 297 MVNPTTRATLEDVASHHWVNW 317

RESULT 4  
 Q8B2N4 PRELIMINARY; PRT; 639 AA.

AC Q8B2N4  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Mus musculus adult male dienecephalon cDNA, RIKEN full-length enriched  
 DE library, clone:9330154N24 product:weakly similar to PROBABLE  
 DE SRINE/THRONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.-).  
 GN Name=1200013B22Rik;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.,  
 RA "High-efficiency full-length cDNA cloning";  
 RL Meth. Enzymol. 303:19-44 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;  
 RX STRAIN=C57BL/6J; TISSUE=Diencephalon;  
 RX MEDLINE=21085660; PubMed=11217851;

RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs";  
 RL Nature 420:563-573 (2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Iashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Hashizawa T.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Iihikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kogawa H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito K., Satoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 DR EMBL; AK034082; BAC28575.1; -;  
 DR HSP; P31751; IGZK.  
 DR MGD; MG1:1921387; 1200013B22R1K.  
 DR GO; GO:0005524; P:ATP binding; IEA.  
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0016740; P:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR011009; Kinase\_like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 639 AA; 70632 MW; FB98EE915C95FA5D CRC64;

Query Match 99.0%; Score 1328; DB 2; Length 639;  
 Best Local Similarity 96.9%; Pred. No. 1.6e-96;  
 Matches 253; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 YEFLETGKGTGKVKVKARESSRLVAIKSIRKDKIKDQDILLHRRREIEIMSSLNHPHI 60  
 DB 57 YEFLETGKGTGKVKVKARESSRLVAIKSIRKDKIKDQDILLHRRREIEIMSSLNHPHI 116  
 QY 61 IAIHE-----VFENSSKIVIVVEYASRGDLVDYISERPRLSERDARHFRQIVSALH 112  
 DB 117 IAIHEVGRSLVTVPENSSKIVIVVEYASRGDLVDYISERPRLSERDARHFRQIVSALH 176  
 QY 113 YCHQNGIVHRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTFGSGPLYASPEIVNGK 172  
 DB 177 YCHQNGIVHRDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTFGSGPLYASPEIVNGK 236  
 QY 173 PYVGPVDSWSLGVLLIYLVHGTMPDGDQDHTLVKQISNGAYREPDKPSDACGLLRWLL 232  
 DB 237 PYVGPVDSWSLGVLLIYLVHGTMPDGDQDHTLVKQISNGAYREPDKPSDACGLLRWLL 296  
 QY 233 MVNPTRRATLIEDVASHWVNW 253  
 DB 297 MVNPTRRATLIEDVASHWVNW 317  
 RESULT 5  
 Q9H093 PRELIMINARY; PRT; 628 AA.  
 AC Q9H093;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DE 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein DKF2p434J037 (Hypothetical protein SNARK)  
 DE (Hypothetical protein FLJ90349).  
 GN Name=DKF2p434J037; Synonyms=SNARK;  
 OS Homo sapiens (Human);  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RA Koehrer K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RC TISSUE=Lymph;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Whiting M., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RC SEQUENCE FROM N.A.  
 RC TISSUE=Lymph;  
 RA Strausberg R.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RC SEQUENCE FROM N.A.  
 RP Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,

RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,  
 RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,  
 RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.,  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 DR EMBL; AL136891; CAB66825.1; -;  
 DR EMBL; BC017306; AAH17306.1; -;  
 DR EMBL; AK074830; BAC11234.1; -;  
 DR HSP; P31751; IGZK.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR011009; Kinase like.  
 DR InterPro; IPR000719; Prot.kinase.  
 DR InterPro; IPR002290; Ser\_Thr\_kinase.  
 DR InterPro; IPR008271; Ser\_Thr\_kinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot.kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Hypothetical protein; Kinase;  
 KW Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 628 AA; 69611 MW; F76F8B1BF94F4C87 CRC64;  
 Query Match 96.3%; Score 1293; DB 2; Length 628;  
 Best Local Similarity 95.7%; Pred. No. 9.3e-94;  
 Matches 242; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 YEFLETGKGTGKGVKKARESGRLVAIKRKDKIKDEQDLLHRRREIEMSSLNPHI 60  
 DB 53 YEFLETGKGTGKGVKKARESGRLVAIKRKDKIKDEQDLLHRRREIEMSSLNPHI 112  
 QY 61 IAIHEVFENSSKIVIMEYASRGDLVDYISERPLSERDARHFRQIVSALHYCHONGIV 120  
 DB 113 IAIHEVFENSSKIVIMEYASRGDLVDYISERQLSREARHFRQIVSAVHYCHQNRVV 172  
 QY 121 HRDLKLENILLDANGNIKIADFGLSNLYHKGFQFTFCGSPLYASPEIVNGKPYVGPEVD 180  
 DB 173 HRDLKLENILLDANGNIKIADFGLSNLYHKGFQFTFCGSPLYASPEIVNGKPYTGPEVD 232  
 QY 181 SWSLGLVLLYLIVHGTMPDGDQDHTLVKQISNGAYRPPKPSDAGLIRLWLLMVPTRRA 240  
 DB 233 SWSLGLVLLYLIVHGTMPDGDHDKILVKQISNGAYRPPKPSDAGLIRLWLLMVPTRRA 292  
 QY 241 TLEDVASHWWVNW 253  
 DB 293 TLEDVASHWWVNW 305  
 RESULT 6  
 ID ARK5 HUMAN STANDARD; PRT; 661 AA.  
 AC O60285;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 05-JUL-2004 (Rel. 44, Last annotation update)  
 GN AMPK-related protein kinase 5 (EC 2.7.1.37).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98290545; PubMed=9628581;  
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,  
 RA Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. IX.

The complete sequences of 100 new cDNA clones from brain which can  
 code for large proteins in vitro.";  
 DNA Res. 5:31-39(1998).  
 (2)  
 RP CHARACTERIZATION, AND MUTAGENESIS OF SER-600.  
 RX MEDLINE=22393479; PubMed=12409306; DOI=10.1074/jbc.M206025200;  
 RA Suzuki A., Kusakai G.-I., Kishimoto A., Lu J., Ogura T., Lavin M.F.,  
 RA Esumi H.;  
 RT "Identification of a novel protein kinase mediating Akt survival  
 signaling to the ATM protein.";  
 RL J. Biol. Chem. 278:48-53(2003).  
 CC -!- FUNCTION: Involved in tolerance to glucose starvation.  
 CC Phosphorylates ATM.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- ENZYME REGULATION: Activated by PKB/AKT1 during glucose  
 starvation.  
 CC -!- TISSUE SPECIFICITY: Expressed at high levels in heart and brain,  
 and at lower levels in skeletal muscle, kidney, ovary, placenta,  
 lung and liver.  
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB011109; BAA25463.1; -;  
 DR HSP; P31751; IGZK.  
 DR MIW; 608130; -;  
 DR InterPro; IPR011009; Kinase like.  
 DR InterPro; IPR000719; Prot.kinase.  
 DR InterPro; IPR008271; Ser\_Thr\_kinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot.kinase; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Phosphorylation; Polymorphism;  
 KW Serine/threonine-protein kinase; Transferase.  
 FT DOMAIN 55 306  
 FT NP\_BIND 61 69 ATP (By similarity).  
 FT ACT\_SITE 178 178 Proton acceptor (By similarity).  
 FT BINDING 84 84 ATP (By similarity).  
 FT MOD\_RES 600 600 Phosphoserine (by PKB/AKT1).  
 FT VARIANT 543 543 P -> R (in dbSNP:3741883).  
 FT FTID=VAR\_017246.  
 FT MUTAGEN 600 600 S->A: No phosphorylation.  
 SQ SEQUENCE 661 AA; 74304 MW; 806F37D52CA4718F CRC64;  
 Query Match 84.1%; Score 1128.5; DB 1; Length 661;  
 Best Local Similarity 82.3%; Pred. No. 1e-80;  
 Matches 209; Conservative 24; Mismatches 20; Indels 1; Gaps 1;  
 QY 1 YEFLETGKGTGKGVKKARESGRLVAIKRKDKIKDEQDLLHRRREIEMSSLNPHI 59  
 DB 55 YEFLETGKGTGKGVKKARESGRLVAIKRKDKIKDEQDMVHRRREIEMSSLNPHI 114  
 QY 60 IAIHEVFENSSKIVIMEYASRGDLVDYISERPLSERDARHFRQIVSALHYCHONGI 119  
 DB 115 IISIEYFENKDKIIVIMEYASKGELYDISSRRLLSERETRHFRQIVSAVHYCHNGV 174  
 QY 120 VHRDLKLENILLDANGNIKIADFGLSNLYHKGFQFTFCGSPLYASPEIVNGKPYVGPEV 179  
 DB 175 VHRDLKLENILLDANGNIKIADFGLSNLYHKGFQFTFCGSPLYASPEIVNGRYPGPEV 234  
 QY 180 DSWSLGLVLLYLIVHGTMPDGDQDHTLVKQISNGAYRPPKPSDAGLIRLWLLMVPTRR 239  
 DB 235 DSWALGVLVLLYVGTMPDGDHDKILVKQISNGAYRPPKPSDAGLIRLWLLMVPDRR 294  
 QY 240 ATLEDVASHWWVNW 253











RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*."  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champagne M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*  
RT melanogaster euchromatic genome sequence."  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk K.S.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review."  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the *Drosophila melanogaster* euchromatic  
RT a genomes perspective."  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003686; AA65135.1; -;  
SQ SEQUENCE 1180 AA; 127743 MW; A88A012664046EC6 CRC64;  
  
Query Match 61.3%; Score 822.5; DB 2; Length 1180;  
Best Local Similarity 61.7%; Pred. No. 3.3e-56;  
Matches 156; Conservative 44; Mismatches 52; Indels 1; Gaps 1;  
  
QY 1 YFLETIGKGYGVKKKA-RESSGLVAIKRKDKIQEDLLHIREIEMSLNHPH 59  
DB 70 FDIITKLGQGYGVKQVGINKETQGEVAIKRKCKIEADLVIRREVOIMSSVHPN 129  
QY 60 IIAIHEVFENSSKIVIVMEYASRGDLYDYISERPLSERDARHFFROIVSALHYCHONGI 119  
DB 130 IIAIHYEFENREKVLWMEFAAGELYDLSERKVLTEEEARRIFRQVATVYCHHKI 189  
QY 120 VHRDLKENILLDANGNIKIADFGSLNLYHKGFLOTFCGSPLYASPEIYNGKPYGVPEV 179  
DB 190 CHRDLKENILLDEKNAKIADFGLSNVFDDQRLLTGFCGSPLYASPEIYVEGTPYQGVPE 249  
QY 180 DWSLSGLVLLTLYVHTGTPFQGDHKTLYVKQISNGAYEPKPSDACGLIRWLLMVNTRR 239  
DB 250 DCWSLSGLVLLTLYVGSMPFSDGNPKRLVKQLSQISQDGYEPRKPSRASTLIRDLMTVCPRK 309

QY 240 ATLEDVASHWVN 252  
DB 310 ASIEQICSHWVN 322  
  
RESULT 13  
QYVH05 PRELIMINARY; PRT; 1427 AA.  
AC Q9VH05; Q9VH04;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DE CG11870-PA (CG11870-pb).  
OS ORFNames=CG11870;  
GN *Drosophila melanogaster* (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; *Drosophila*.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Flesler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*."  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champagne M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*  
RT melanogaster euchromatic genome sequence."  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirska R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomic perspective";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaninkur J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RX SEQUENCE FROM N.A.
RG FLYBASE;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RX FLYBASE;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase families.
DR EMBL; AE003686; AAF54517.3; --
DR HSP; O63450.1A06.
DR InAct; Q9VH05; --
DR FlyBase; FBgn0037804; CG11870.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1427 AA; 153443 MW; 530928D0A057B5C8 CRC64;

Query Match 61.3%; Score 822.5; DB 2; Length 1427;
Best Local Similarity 61.7%; Pred. NO. 4.1e-56;
Matches 156; Conservative 44; Mismatches 52; Indels 1; Gaps 1;

QY 1 YEFLETGKGYGVKKKA-RESSRLVAIKRKDKIKDEODLHRIEIMSSLNHPH 59
DB 70 FDIINKGQGYGVQLGINKETGEVAIKTKICKIEADLVIRREVQIMSVVHPN 129

QY 60 IIAIHEVPENSKIVIMVEASRGDLYDIISERPLSERDARHFFRQIVSALHYCHQNGI 119
DB 130 ITHIVFENREKMWLVNVEAAGGELYLSERKVLVTEEARIRFRQVATAVYCHKHKI 189

QY 120 VHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPIVNGKPYGVPEV 179
DB 190 CHRDLKLENILLDEKGNAKIADFGLSNVDFDQRLGLTFCGSPLYASPIVGTQYQGEV 249

QY 180 DWSLGLVLLYILVHGTMPFDGODHKTLLVKQISNGAYRPPKPSDACGLIRLLWVNPTR 239
DB 250 DWSLGLVLLYILVYSGMFDGNSFKRLVKQISQGDYIEPRKESASTLIIRDMLIVCPKR 309

QY 240 ATLEDVASHHWN 252

```

```

Db 310 ASIEQICSHWVN 322
RESULT 14
AAF54517 PRELIMINARY; PRT; 1427 AA.
AC AAF54517.
DT 01-APR-2004 (TREMBlrel. 27, Created)
DT 01-APR-2004 (TREMBlrel. 27, Last sequence update)
DT 01-APR-2004 (TREMBlrel. 27, Last annotation update)
DE CG11870-PA.
GN CG11870.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hestlin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhao M., Zhang G., Zhou Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacieb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirska R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Wheeler S.E., Myers S.E., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.

```



Image Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 11:04:39 ; Search time 5813.23 Seconds  
(without alignments)  
11866.113 Million cell updates/sec

Title: US-09-980-464-4\_COPY\_123\_2015

Perfect score: 1893

Sequence: 1 atggagtcggtggccttact.....gaatctgctcaagctcagc 1893

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1.\*

2: gb\_est2.\*

3: gb\_hic.\*

4: gb\_est3.\*

5: gb\_est4.\*

6: gb\_est5.\*

7: gb\_est6.\*

8: gb\_gss1.\*

9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1893	100.0	2899	3 AK004737	AK004737 Mus muscu
2	1893	100.0	3098	3 AK033672	AK033672 Mus muscu
3	1857.4	98.1	2869	3 AK034082	AK034082 Mus muscu
4	1773	93.7	1896	9 AY399041	AY399041 Mus muscu
5	1261	66.6	1887	9 AY399039	AY399039 Homo sapi
6	948.4	50.1	1887	9 AY399040	AY399040 Pan trogl
7	683.6	36.1	726	6 CB248251	CB248251 UI-M-FD0-
8	672	35.5	672	7 CO424322	CO424322 UI-M-HU0-
9	652.4	34.5	655	7 CK781308	CK781308 UI-M-GI0-
10	651.6	34.4	805	4 BG174288	BG174288 602334589
11	646	34.1	646	4 BG694881	BG694881 NISC_iv09
12	645.8	34.1	948	2 BF789245	BF789245 602105144
13	640	33.8	795	6 CA319312	CA319312 UI-M-FW0-
14	627.6	33.2	782	7 CN261003	CN261003 170004247
15	614.8	32.5	1069	4 BM927376	BM927376 AGENCOURT
16	605.8	32.0	1022	5 BQ062868	BQ062868 AGENCOURT
17	603	31.9	893	2 BF785290	BF785290 602108490
18	596	31.5	705	7 CF726196	CF726196 UI-M-GZ0-
19	586.4	31.0	1005	5 BQ060729	BQ060729 AGENCOURT
20	596.4	31.0	1029	5 BQ059508	BQ059508 AGENCOURT
21	578	30.5	722	4 B1143493	B1143493 602907673
22	574.4	30.3	1012	5 BQ064231	BQ064231 AGENCOURT
23	562.8	29.7	713	2 BB625283	BB625283 BB625283
24	558.8	29.5	667	4 B1648260	B1648260 603278202

25	554.8	29.3	558	8 BH116351	BH116351 RPCI-24-2
26	552.2	29.2	1042	4 BM909401	BM909401 AGENCOURT
27	547	28.9	939	2 BE912458	BE912458 601666074
28	541.4	28.6	761	6 CD632904	CD632904 55046560H
29	541.4	28.6	851	5 BU228581	BU228581 603947039
30	535.2	28.3	824	7 CK483217	CK483217 AGENCOURT
31	534	28.2	1003	5 BQ056806	BQ056806 AGENCOURT
32	532.8	28.1	890	6 CA487493	CA487493 AGENCOURT
33	531.8	28.1	535	4 BG086714	BG086714 H3129F11-
34	530	28.0	1120	4 BM909505	BM909505 AGENCOURT
35	529.2	28.0	760	4 EG753131	EG753131 602732207
36	528	27.9	1038	5 BQ056633	BQ056633 AGENCOURT
37	526.2	27.8	1047	5 BQ064127	BQ064127 AGENCOURT
38	524.8	27.7	865	4 BG172224	BG172224 602333366
39	522.4	27.6	616	4 B1647932	B1647932 603276014
40	521	27.5	818	7 CK483904	CK483904 AGENCOURT
41	508	26.8	556	6 CA535696	CA535696 C0222H05-
42	506.4	26.8	1044	5 BQ065997	BQ065997 AGENCOURT
43	505.8	26.7	699	6 CD632908	CD632908 55053832H
44	498.4	26.3	1086	5 BQ056490	BQ056490 AGENCOURT
45	495.8	26.2	1034	5 BQ055919	BQ055919 AGENCOURT

#### ALIGNMENTS

RESULT 1	AK004737	2899 bp	linear	HTC 03-APR-2004
LOCUS	AK004737	Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200013B22 product:weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.1-) [Homo sapiens], full insert sequence.	mrna	
DEFINITION	AK004737	GI:26334437		
ACCESSION	AK004737	HTC; CAP trapper.		
VERSION	AK004737.2	Mus musculus (house mouse)		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus			
ORGANISM	Mus musculus			
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			
REFERENCE	3			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, Y., Yoneeda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer			
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)			
MEDLINE	20530913			
PUBMED	11076861			
REFERENCE	4			
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.			
TITLE	Functional annotation of a full-length mouse cDNA collection			
JOURNAL	Nature 409, 685-690 (2001)			
REFERENCE	5			

**AUTHORS**  
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

**TITLE**  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

**JOURNAL REFERENCE**  
Nature 420, 563-573 (2002)

**AUTHORS**  
6 (bases 1 to 2899)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hangaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

**TITLE**  
Direct Submission

**JOURNAL**  
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

**COMMENT**  
On Dec 10, 2002 this sequence version replaced gi:12836134. Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

**FEATURES**  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGCGCGCACTCGAGTCTTTTCTTTTNN 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAGCGCGCACTCGAGTCTTTTCTTTTNN 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOUR.

**source**  
Location/Qualifiers  
1..2899  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:1200013B22"  
/db\_xref="taxon:10090"  
/clone="1200013B22"  
/sex="male"  
/tissue\_type="lung"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
94..1989

**CDS**  
/notes="unnamed protein product; putative weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0517 (EC 2.7.1.-) (Homo sapiens) (SWISSPROT|060285, evidence: PASTY, 56.8%id, 98.6%length, match=1849)";  
/codon\_start=1  
/protein\_id="BAB23518.1"  
/db\_xref="GI:12836135"  
/translations="MESVALLQRPQAPSALASARPLADGLIKSPKPLMKQAV KRHHKHLNLRHVEFTLGKTVGVKKARESSGRVLAKSRKDKIDQDOLLHIRETEIMSSLNPHIIAIEVFENSSKIVIVWEYASRGDLIDYISERPRISERDARHFRQIVSALVHQNGQIVHRDUDKLEILLDANGNIKIADFGSLNDLHNGKFLQTPCGSPLYASPEIVNGQYGVPEVDSVLSGLYILVHGTMPDFGQDHKTLVLIQINAYRBPBKPSDACGLIIRWLLMVFTRATLVDASHVMVNNVYTTGVGEALREGHPSGDFQFRA SMADWLRSSRPKSLKPKGILKKSSTSSGEVQDQELRPVPTPCQVPVAVSILGDAEDTSSRPKSLKPKGILKKSSTSSGEVQDQELRPVPTPCQVPVAVSILPRKGIKKSQPSGYSVSPERSESGELLDASVDFVSGDVEQKSFQAGLLHRRGI LKUNGKFSRTALSGTFTSPFSLDQLASHPAARPSRPSGAVSEDSILSESFDQLDPERLPFTPLRGCVSDNLRLEQPPSEGLKRWQESLFGSDSCFSLTDCQEVTAAYRALGICSKLS"

**polyA\_signal** 2876..2881 /notes="putative"  
**polyA\_site** 2899 /notes="putative"  
**ORIGIN**

Query Match	100.0%;	Score 1893;	DB 3;	Length 2899;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1893;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

QY 1 ATCGAGTCCGCTGCGCTTACTCCAGCGCCGAGCAGAGCTCCCTCGGCTCGCGCTCGGCC 60
    |||
Db 94 ATCGAGTCCGCTGCGCTTACTCCAGCGCCGAGCAGAGCTCCCTCGGCTCGCGCTCGGCC 153
    |||
QY 61 TCGGAGAGCGCCCGCGCGCTGCGGACGGGCTCATCAAGTCGCTTAACTCTGATGAAG 120
    |||
Db 154 TCGGAGAGCGCCCGCGCGCTGCGGACGGGCTCATCAAGTCGCTTAACTCTGATGAAG 213
    |||
QY 121 AAGCAGCGGTGAAGCGGCACCATCAACACAACTCGGCGCAGCTACGAGTTCCTG 180
    |||
Db 214 AAGCAGCGGTGAAGCGGCACCATCAACACAACTCGGCGCAGCTACGAGTTCCTG 273
    |||
QY 181 GAGACGCTGGGCAAGGGCACCTACGGGAAGGTCAAGAAGGCACGAGAGCTCGGGCGT 240
    |||
Db 274 GAGACGCTGGGCAAGGGCACCTACGGGAAGGTCAAGAAGGCACGAGAGCTCGGGCGT 333
    |||
QY 241 CTGGTGCCATCAAGTTCATCAGGAAGACAAAATCAAGATGAGCAGGATCTGTGCAC 300
    |||
Db 334 CTGGTGCCATCAAGTTCATCAGGAAGACAAAATCAAGATGAGCAGGATCTGTGCAC 393
    |||
QY 301 ATACGAGGAGAGATTCAGATCATGTCTCACTCAACACCCACACATCATTTGCCATCAT 360
    |||
Db 394 ATACGAGGAGAGATTCAGATCATGTCTCACTCAACACCCACACATCATTTGCCATCAT 453
    |||
QY 361 GAAGTGTGAGAAATAGCAGCAAGATTGTGATGTGTCATGAGTATGCGACCGAGGCGAT 420
    |||
Db 454 GAAGTGTGAGAAATAGCAGCAAGATTGTGATGTGTCATGAGTATGCGACCGAGGCGAT 513
    |||
QY 421 CTGTATGATTACATCAGTAGCGGCGCACCGCTGAGTGAGCGGACCGCAGGCAATTTCTTC 480
    |||
Db 514 CTGTATGATTACATCAGTAGCGGCGCACCGCTGAGTGAGCGGACCGCAGGCAATTTCTTC 573
    |||
QY 481 CGACAGATCGTGTCTCCCTGCACTACTGCACACAGAAAGCGGATCGTTCAACCGAGATCTC 540
    |||
Db 574 CGACAGATCGTGTCTCCCTGCACTACTGCACACAGAAAGCGGATCGTTCAACCGAGATCTC 633
    |||
QY 541 AAGCTGGAACAACTCTTTCTAGATGCAATGGAACATCAAGATTGCTGATTTGGCCTC 600
    |||
Db 634 AAGCTGGAACAACTCTTTCTAGATGCAATGGAACATCAAGATTGCTGATTTGGCCTC 693
    |||
QY 601 TCCAACTGTACCAAAAGCAAGTTCTCTCCAGACGTTCTGTGGGAGCCCTCTCTAGGCC 660
    |||
Db 694 TCCAACTGTACCAAAAGCAAGTTCTCTCCAGACGTTCTGTGGGAGCCCTCTCTAGGCC 753
    |||
QY 661 TCGCCTTGAGATAGTCAACGGGAAGCCCTATGTGGGCGCAGAGGTGACAGCTGCTCTCTG 720
    |||
Db 754 TCGCCTTGAGATAGTCAACGGGAAGCCCTATGTGGGCGCAGAGGTGACAGCTGCTCTCTG 813
    |||
QY 721 GCGTGTCTCTGTACATCTCTGTTGATGCGACCATCCCTTTTACCGGGCAGGATCATAAA 780
    |||
Db 814 GCGTGTCTCTGTACATCTCTGTTGATGCGACCATCCCTTTTACCGGGCAGGATCATAAA 873
    |||
QY 781 ACACCTGTTGAGCAAAATCAGTAAACGGGCTTACCGTGAGCCGCCCAAGCGCTCCGATGCC 840
    |||
Db 874 ACACCTGTTGAGCAAAATCAGTAAACGGGCTTACCGTGAGCCGCCCAAGCGCTCCGATGCC 933
    |||
QY 841 TGTGGCCTCATCCGGTGGCTGTTTAAATGTTGAAACCCACCCCGTGGGCCCACTCTGGAGGAT 900
    |||
Db 934 TGTGGCCTCATCCGGTGGCTGTTTAAATGTTGAAACCCACCCCGTGGGCCCACTCTGGAGGAT 993
    |||
QY 901 GTAGCCAGTCAATTGGTGGGTCAACTGGGGTTACACACCGGAGTCCGGGAAAGAGGAGCC 960
    |||
Db 994 GTAGCCAGTCAATTGGTGGGTCAACTGGGGTTACACACCGGAGTCCGGGAAAGAGGAGCC 1053
    |||

```



Qy	961	CTGCGTGGGTCACCCCTAGTGTGCACTTTGGCCGGGCTCCATGGCGGACTGGTTA	1020
Db	1054	CTGCGTGGGTCACCCCTAGTGTGCACTTTGGCCGGGCTCCATGGCGGACTGGTTA	1113
Qy	1021	CGTGCCTCTCGCCGCCCTCTCGGAGAAATGGAGCCAAAGTGTGCGAGCTTCTTCAAGCAG	1080
Db	1114	CGTGCCTCTCGCCGCCCTCTCGGAGAAATGGAGCCAAAGTGTGCGAGCTTCTTCAAGCAG	1173
Qy	1081	CACGTGCGGAGGTGAAGCACTGTACCTGGGCTGGAGCGGCAACATCTCTTAAAGAAG	1140
Db	1174	CACGTGCGGAGGTGAAGCACTGTACCTGGGCTGGAGCGGCAACATCTCTTAAAGAAG	1233
Qy	1141	TCCGAAAGGAGATGATGGCTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCTCT	1200
Db	1234	TCCGAAAGGAGATGATGGCTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCTCT	1293
Qy	1201	TCTGCCCTCGCAAGAGAGCGCTTAAGCTTCCGAAAGGCATTTCTCAAGAAAGTCTCTCT	1260
Db	1294	TCTGCCCTCGCAAGAGAGCGCTTAAGCTTCCGAAAGGCATTTCTCAAGAAAGTCTCTCT	1353
Qy	1261	ACCTCGTCAGGGAGGTACAGGAGGACCTCAGGAATCTCAGACCGGTGCCTGATCTCCA	1320
Db	1354	ACCTCGTCAGGGAGGTACAGGAGGACCTCAGGAATCTCAGACCGGTGCCTGATCTCCA	1413
Qy	1321	GGGCGAGCTGTCTCCTGTGTATCCTCTCCCAAGGAAAGGCATCTTAAAGAAAGTCTCGA	1380
Db	1414	GGGCGAGCTGTCTCCTGTGTATCCTCTCCCAAGGAAAGGCATCTTAAAGAAAGTCTCGA	1473
Qy	1381	CAGGTGAATCTGTTACTACTCTCTCCAGAGCCAGAGTCTGGGGAACTCTTAGAC	1440
Db	1474	CAGGTGAATCTGTTACTACTCTCTCCAGAGCCAGAGTCTGGGGAACTCTTAGAC	1533
Qy	1441	GCAGTGAATCTGTTGAGTGGGACCCCGTGGAGCAGAGTCTCCACAGGCTTCAGGG	1500
Db	1534	GCAGTGAATCTGTTGAGTGGGACCCCGTGGAGCAGAGTCTCCACAGGCTTCAGGG	1593
Qy	1501	CTCTCTCTCCACCGCAAGGCGATTTCTCAAACTCAATGGCAAGTTCTCCCGCACAGCCTTA	1560
Db	1594	CTCTCTCTCCACCGCAAGGCGATTTCTCAAACTCAATGGCAAGTTCTCCCGCACAGCCTTA	1653
Qy	1561	GAGGCACTACCTAGACCTTTGGCTCCCTGAGCAACCTGGCTCTCTCCCATCTCGCA	1620
Db	1654	GAGGCACTACCTAGACCTTTGGCTCCCTGAGCAACCTGGCTCTCTCCCATCTCGCA	1713
Qy	1621	GCCGGCCAGCCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTCTGCTCCGAGTCC	1680
Db	1714	GCCGGCCAGCCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTCTGCTCCGAGTCC	1773
Qy	1681	TTTGACCAATTGGACTTGGCTGACGCTTTCCCGGAAACCCCACTGAGGGGCTGTGTGCT	1740
Db	1774	TTTGACCAATTGGACTTGGCTGACGCTTTCCCGGAAACCCCACTGAGGGGCTGTGTGCT	1833
Qy	1741	GTGGAACCTAGGGGGCTTGGAGCAGCTCCCTCAGAAAGTCTGAAGCGATGTGGCAG	1800
Db	1834	GTGGAACCTAGGGGGCTTGGAGCAGCTCCCTCAGAAAGTCTGAAGCGATGTGGCAG	1893
Qy	1801	GAATCTTGGGGATAGCTGCTTTCTGACAGACTGCCAAGAGGTGACTGCGAGCCTAC	1860
Db	1894	GAATCTTGGGGATAGCTGCTTTCTGACAGACTGCCAAGAGGTGACTGCGAGCCTAC	1953
Qy	1861	AGACAAGCCCTAGGAATCTGCTCAAAGCTCAGC	1893
Db	1954	AGACAAGCCCTAGGAATCTGCTCAAAGCTCAGC	1986

## RESULT 2

AK033672  
LOCUS  
DEFINITION  
Mus musculus adult male cecum cDNA, RIKEN full-length enriched library, clone:9130215K18 product:weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.-) [Homo sapiens], full insert sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AK033672  
GI:26329364  
HTC: CAP trapper.  
Mus musculus (house mouse)

## REFERENCE

1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636

## REFERENCE

2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

## REFERENCE

3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

## REFERENCE

4  
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

## REFERENCE

5  
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 3098)

## REFERENCE

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission

## REFERENCE

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/  
Location/Qualifiers  
1. .3098  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM.DB:9130215K18"  
/db\_xref="taxon:10090"  
/clone="9130215K18"  
/sex="male"  
/tissue type="cecum"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
93. .1988  
/note="unnamed protein product; putative weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE KIA0537 (EC 2.7.1.-) [Homo sapiens] (SWISSPROT|O60285, evidence: FASTY, 56.8%ID, 98.6%length, match=1849)"  
/codon\_start=1  
/protein\_id="BAC28421.1"  
/db\_xref="GI:26329365"  
/translation="MESVALLQRPSPASALASERPLADGLIKSPKPLMKQAV KRHHKNSLRHRYFLTLGKTYGKVKARESSRLVAIKSRDKIKDEQDILLHR REIEMSLNHPHIIAHEVPENSSKIVIMEVASRGDLVDYISERPLSERDARHP ROIVSALHYCHQNGIHRDLKLENILLDANGNIKIADGLSNLYHKGFLOTGCSPL YASPEIYNGPYVGPEDVSNLSGLVLLYLVIHGTMPDQDHLTKVKSINGAREPK SMADGLIRLLMWNPRATILEDVASHWNWNGYITTVGVEQEARLREGGHSPGDFRA SMDLRRSRPLLENGACVSPFKQHPVGGSTVPLGRQHSLSKRKENDMAQNLQ GDAEDTSRPSKSLKPKGILKKKSSTSEGEVQEDPQLRPVDPDTPGQFPAVSLI PRKGIILKRSQRSGSYSPSPESGELLDAISDFVSGDPVEQKSPOASGLLLHRKGI LKNGKFSRTALEGTTFTSGSLDOLASSHPARPSPGSAVSEDSITLSESPDQDL PERLPETPLRGCVSDNLRLEQPPSGLKRWQESLGDSCFSITDQCVTAAYRQAL GICSKLS"  
3082. .3087  
/note="putative"  
3098  
/note="putative"

polyA\_signal  
polyA\_site

ORIGIN

Query Match 100.0%; Score 1893; DB 3; Length 3098;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGTCCGTGGCTTACTCCAGCGCCGAGCAGGCTCCCTCGGCTCCGCTGGCC 60  
DB 93 ATGAGTCCGTGGCTTACTCCAGCGCCGAGCAGGCTCCCTCGGCTCCGCTGGCC 152  
QY 61 TCGAGAGCGCCGCGCTGGCGGACGGCTCATCAAGTCCGCTAAACCTCTGATGAAG 120  
DB 153 TCGAGAGCGCCGCGCTGGCGGACGGCTCATCAAGTCCGCTAAACCTCTGATGAAG 212  
QY 121 AAGCAGCGGTGAAGCGGACCATCAAAACACACACACCTGGGACCGCTAGGATTCCTG 180  
DB 213 AAGCAGCGGTGAAGCGGACCATCAAAACACACACACCTGGGACCGCTAGGATTCCTG 272  
QY 181 GAGACGCTGGCGGACGACCTACGGAAGGTGAAGAGGACGAGAGAGCTCGGGCGT 240  
DB 273 GAGACGCTGGCGGACGACCTACGGAAGGTGAAGAGGACGAGAGAGCTCGGGCGT 332  
QY 241 CTGGTGCCCATCAAGTCCATCAGAAAGACAAAATCAAGATGAGCAGGATCTGCTGCAC 300  
DB 333 CTGGTGCCCATCAAGTCCATCAGAAAGACAAAATCAAGATGAGCAGGATCTGCTGCAC 392  
QY 301 ATACGAGGAGAGATTGAGATCATGTCTTCACTCAACACACCCACATCATTCGCATCCAT 360  
DB 393 ATACGAGGAGAGATTGAGATCATGTCTTCACTCAACACACCCACATCATTCGCATCCAT 452  
QY 361 GAAGTGTGAGAAATAGCAGCAAGATTGTGATGTGATGAGATGCGACCGAGCGGAT 420  
DB 453 GAAGTGTGAGAAATAGCAGCAAGATTGTGATGTGATGAGATGCGACCGAGCGGAT 512  
QY 421 CTGTATGATTACATCAGTGAAGCGGACCGCTGAGTGAAGCGGACCGCAGGCAATTCCTTC 480

DB 513 CTGTATGATTACATCAGTGAAGCGGACCGCTGAGTGAAGCGGACCGCAGGCAATTCCTTC 572  
QY 481 CGACAGATCGTGTGTCGCCCTGCACTACTGCGCACAGAACCGGATCGTTCACCGAGATCTC 540  
DB 573 CGACAGATCGTGTGTCGCCCTGCACTACTGCGCACAGAACCGGATCGTTCACCGAGATCTC 632  
QY 541 AAGCTGGAACAAATCCTTCTAGATGCAATGGAACAAATCAAGATTGCTGATCTTGGCTC 600  
DB 633 AAGCTGGAACAAATCCTTCTAGATGCAATGGAACAAATCAAGATTGCTGATCTTGGCTC 692  
QY 601 TCAAACTGTACCAACAAAGCAAGTTCCTCCAGACGTTCTGTGGAGCCCTCTCTACGCC 660  
DB 693 TCAAACTGTACCAACAAAGCAAGTTCCTCCAGACGTTCTGTGGAGCCCTCTCTACGCC 752  
QY 661 TCGCCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCAGAGGTGGAACAGTCTGCTCTG 720  
DB 753 TCGCCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCAGAGGTGGAACAGTCTGCTCTG 812  
QY 721 GCGTTCCTCTGTACATCTCTGGTGCATGGCACCATGCGCTTTGACGGGCAAGGATCAAAA 780  
DB 813 GCGTTCCTCTGTACATCTCTGGTGCATGGCACCATGCGCTTTGACGGGCAAGGATCAAAA 872  
QY 781 ACATGCTGAAGCAAAATCAGTAACGGGCTTACCGTGAGCGGCCCAAGCCGTCGATGCC 840  
DB 873 ACATGCTGAAGCAAAATCAGTAACGGGCTTACCGTGAGCGGCCCAAGCCGTCGATGCC 932  
QY 841 TGTGGCTGATCCGCTGGCTGTTAAATGTTGAAACCCACCCGTCGGGCCACACATGGAGGAT 900  
DB 933 TGTGGCTGATCCGCTGGCTGTTAAATGTTGAAACCCACCCGTCGGGCCACACATGGAGGAT 992  
QY 901 GTAGCAGATCATGTTGGTCACTGGGTTTACACACCGGAGTCGGGGAACAGGAGCC 960  
DB 993 GTAGCAGATCATGTTGGTCACTGGGTTTACACACCGGAGTCGGGGAACAGGAGCC 1052  
QY 961 CTGCTGAGGCTGGGACCCCTAGTGTGATCTTGGCGGGCCCTCCATGGCGGACTGTTA 1020  
DB 1053 CTGCTGAGGCTGGGACCCCTAGTGTGATCTTGGCGGGCCCTCCATGGCGGACTGTTA 1112  
QY 1021 CGTGGCTCTCGGCCCTCTCTGGAGAAATGGAAGCAAGGTGTGAGCTTCTTCAAGCAG 1080  
DB 1113 CGTGGCTCTCGGCCCTCTCTGGAGAAATGGAAGCAAGGTGTGAGCTTCTTCAAGCAG 1172  
QY 1081 CAGTCCGGGAGTGGAAAGCACTGTACCTGGCTGAGCGGCAACATCTCTTAAAGAG 1140  
DB 1173 CAGTCCGGGAGTGGAAAGCACTGTACCTGGCTGAGCGGCAACATCTCTTAAAGAG 1232  
QY 1141 TCCGAAAGGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCTCT 1200  
DB 1233 TCCGAAAGGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCTCT 1292  
QY 1201 TCTCGCCCTGGCAAGAGAGCTTAAAGTTCGAAAGGCAATCTCAAGAAAAGTCTCT 1260  
DB 1293 TCTCGCCCTGGCAAGAGAGCTTAAAGTTCGAAAGGCAATCTCAAGAAAAGTCTCTCT 1352  
QY 1261 ACCTCGTCAGGGAGGTACAGAGGACCTCAGAGAACTCAGACCGGTCTGATCTCCA 1320  
DB 1353 ACCTCGTCAGGGAGGTACAGAGGACCTCAGAGAACTCAGACCGGTCTGATCTCCA 1412  
QY 1321 GGGACGCTGTCCCTGCTGTATCCCTGCTCCCAAGAAAAGGCAATCTTAAAGAGTCTGA 1380  
DB 1413 GGGACGCTGTCCCTGCTGTATCCCTGCTCCCAAGAAAAGGCAATCTTAAAGAGTCTGA 1472  
QY 1381 CAGCGTGAATCTGTTTACTTACTCTCTCCAGAGCCAGGAGTCTGGGAACTCTTAGAC 1440  
DB 1473 CAGCGTGAATCTGTTTACTTACTCTCTCCAGAGCCAGGAGTCTGGGAACTCTTAGAC 1532  
QY 1441 GCCAGTGTGTGTTGTGAGTGGGACCCCGTGGAGCAGAAGTCTCCACAGGCTTCAGGG 1500  
DB 1533 GCCAGTGTGTGTTGTGAGTGGGACCCCGTGGAGCAGAAGTCTCCACAGGCTTCAGGG 1592  
QY 1501 CTCCTCTCCACCGCAAGGGCAATCTCAAACTCAATGGAAGTCTTCCCGCACAGCCTTA 1560  
DB 1593 CTCCTCTCCACCGCAAGGGCAATCTCAAACTCAATGGAAGTCTTCCCGCACAGCCTTA 1652

QY	1561	GAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCCTCCTCCCATCTGCA	1620
Db	1653	GAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCCTCCTCCCATCTGCA	1712
QY	1621	GCCCGGCCAGCCGCCCTCAGGGCTGTGAGTGAGGACAGCATCTGTCTCGAGTCC	1680
Db	1713	GCCCGGCCAGCCGCCCTCAGGGCTGTGAGTGAGGACAGCATCTGTCTCGAGTCC	1772
QY	1681	TTTGACCAATTGGACTTGGCTGAACGCTTTCGCCGAACCCCACTGAGGGCTGTGTCT	1740
Db	1773	TTTGACCAATTGGACTTGGCTGAACGCTTTCGCCGAACCCCACTGAGGGCTGTGTCT	1832
QY	1741	GTGACCAACTGAGGGGCTTGAGCAGCCTCCCTCAGAAGGCTCTGAAGCGATGTGGCAG	1800
Db	1833	GTGACCAACTGAGGGGCTTGAGCAGCCTCCCTCAGAAGGCTCTGAAGCGATGTGGCAG	1892
QY	1801	GAATCTTGGGGGATAGTCTCTTCTCAGACAGTCCCAAGAGTGAATGAGCCTTAC	1860
Db	1893	GAATCTTGGGGGATAGTCTCTTCTCAGACAGTCCCAAGAGTGAATGAGCCTTAC	1952
QY	1861	AGACAAGCCCTAGGAATCTCTCAAGCTCAGC	1893
Db	1953	AGACAAGCCCTAGGAATCTCTCAAGCTCAGC	1985
RESULT 3			
LOCUS	AK034082		
DEFINITION	Mus musculus adult male diencephalon cDNA, RIKEN full-length		
	enriched library, clone:9330154N24 product:weakly similar to		
	PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.-)		
	[Homo sapiens], full insert sequence.		
ACCESSION	AK034082		
VERSION	AK034082.1	GI:26329672	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to		
	prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		
	Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M.,		
	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,		
	Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,		
	Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,		
	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,		
	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format		
	sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE	4		
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the		
	PANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409, 685-690 (2001)		
REFERENCE	5		

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## CDS

## ORIGIN

Query Match 98.1%; Score 1857.4; DB 3; Length 2869;

Best Local Similarity 98.7%; Pred. No. 0;

Matches 1892; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY 1 ATGGAGTCGTGGCCTTACTCCAGCGCCGAGCAGGCTCCCTCGGCTCCGCCCTCGCCGCGCC

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 2869)

Adachi, J., Aizawa, K., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@res.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

Location/Qualifiers

1. 2869

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="FANTOM DB:9330154N24"

/db\_xref="taxon:10090"

/clone="9330154N24"

/sex="male"

/tissue type="diencephalon"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="adult"

98. 2017

/notes="unnamed protein product; putative

weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE

KIAA0537 (EC 2.7.1.-) (Homo sapiens) (SWISSPROT|O60285,

evidence: FASTY, 56.8%ID, 98.6%length, match=1849)"

/codon\_start=1

/protein\_id="BAC28575.1"

/db\_xref="GI:26329673"

/translation="MESGALLQRPQSPASALASARPLADGLINKPKPLMKQAV

KRHKKHNRHRYFLETGKTVGKVKARSSRLVAIKSRKDKIDQDQLLHR

RELEIMSSLNPHPIAIHEVGSRLVTFVENSKIIVMEYASRGDLYDYSRPRLS

ERDARFSPQIYVSAHYCHQNGIVHRDLKLENILLDANGNKIADGLSLNLYKGRFL

QFCGSPFVSPENKVGYPGVDSWLVLLYLVHGTMPFDGDDHLLTVKQLSN

GYSPDFKPSDACGLIRLLWLNPTRRATLEDVASHVWVNWGTVTGQGEALREGGH

PSGDFGRASMDWLRSSRPKSLKLPKGLIKKSKTSSEVOEDPOELRPVPTQO

NPAVSLPRKGLKRSRQESGYSSPEPSGELLADSDVPSGDVPEQKSPQASG

LLLRKGLKNGKFSRTALEGTPSTFGSLDQLASSHPAARSRPSGAVSESLSS

ESFDQDLPLRPETPLPGCVSDNLRGLQEPPEBGLKRWQBSLGDSCFSLTDCQEV

TAAYRQALGICSKLS"

Db	98	ATCGAGTCGGGGCTTTACTCCAGCGCCCGAGCCAGGCTCCCTCGGCTCGGCCCTGGCC	157
Qy	61	TCGAGAGCGCCCGGCGCTGGCGGACGGGCTCATCAAGTCGCTAAACCTCTGATGAAG	120
Db	158	TCGAGAGCGCCCGGCGCTGGCGGACGGGCTCATCAAGTCGCTAAACCTCTGATGAAG	217
Qy	121	AAGCAGCGGTTGAAGCGGCAACATCAAAACACAACTCGCGCACCGCTACGAGTTCCCTG	180
Db	218	AAGCAGCGGTTGAAGCGGCAACATCAAAACACAACTCGCGCACCGCTACGAGTTCCCTG	277
Qy	181	GAGACGCTGGCAAGGGCACTCAGGAAAGGTGAAGAGGCACGAGAGACTCGGGCGT	240
Db	278	GAGACGCTGGCAAGGGCACTCAGGAAAGGTGAAGAGGCACGAGAGACTCGGGCGT	337
Qy	241	CTGGTGGCATCAAGTCATCAGAGAAAGACAAATCAAGATGAGCAGGATCTGCTGCAC	300
Db	338	CTGGTGGCATCAAGTCATCAGAGAAAGACAAATCAAGATGAGCAGGATCTGCTGCAC	397
Qy	301	ATACGGAGGGAGATTGAGATCATCTTTCACTCAACCAACCCACATCATTTGCCATCCAT	360
Db	398	ATACGGAGGGAGATTGAGATCATCTTTCACTCAACCAACCCACATCATTTGCCATCCAT	457
Qy	361	GA-----GTGTTTGAGAAATAGCAAGATTTGATTTGTC	396
Db	458	GAAGTGGSCAGATCACGCTTAGTGACTGTGTTTGAGAAATAGCAGCAAGATTGATTTGTC	517
Qy	397	ATGAGATATGCAGCGAGGGCGATCTGTATGATTTACATCAGTAGCGGCACCGCTGAGT	456
Db	518	ATGAGATATGCAGCGAGGGCGATCTGTATGATTTACATCAGTAGCGGCACCGCTGAGT	577
Qy	457	GAGCGGAGCGCAGGCAATTTCTTCGACAGATCGTGTCTGCCCTGCACTACTGCGCACAG	516
Db	578	GAGCGGAGCGCAGGCAATTTCTTCGACAGATCGTGTCTGCCCTGCACTACTGCGCACAG	637
Qy	517	AACGGGATCGTTTCAACCGAGATCTCAAGCTGGAACACATCTTCTTAGATGCCAATGGAAC	576
Db	638	AACGGGATCGTTTCAACCGAGATCTCAAGCTGGAACACATCTTCTTAGATGCCAATGGAAC	697
Qy	577	ATCAAGATATGCTGACTTTGGCTCTCAGACCTGTACCAAGGCAAGTTCTCTCCAGACG	636
Db	698	ATCAAGATATGCTGACTTTGGGCTCTCAGACCTGTACCAAGGCAAGTTCTCTCCAGACG	757
Qy	637	TTCTGTGGAGCCCTCTCTACGCTCGCTCAGATAGTCAACGGGAAGCCCTATGTGGGC	696
Db	758	TTCTGTGGAGCCCTCTCTACGCTCGCTCAGATAGTCAACGGGAAGCCCTATGTGGGC	817
Qy	697	CCAGAGTGGACAGCTGGTCTCTGGGCGCTTCTCTGTATCACTCTGGTGCATGGCACCATG	756
Db	818	CCAGAGTGGACAGCTGGTCTCTGGGCGCTTCTCTGTATCACTCTGGTGCATGGCACCATG	877
Qy	757	CCCTTTGACGGGACGGATCAATAACAATCTGGTGAAGCAATCAGTAAACGGGCTTACCGT	816
Db	878	CCCTTTGACGGGACGGATCAATAACAATCTGGTGAAGCAATCAGTAAACGGGCTTACCGT	937
Qy	817	GAGCGCCCAAGCCGTCGATGCGCTGTGCTGATCCGGTGGCTGTTAATGGTGAACCCC	876
Db	938	GAGCGCCCAAGCCGTCGATGCGCTGTGCTGATCCGGTGGCTGTTAATGGTGAACCCC	997
Qy	877	ACCCGTCGGGCCACATCGAGAGATAGCCAGTCAATTGGTGGGTCAACTGGGGTTACCC	936
Db	998	ACCCGTCGGGCCACATCGAGAGATAGCCAGTCAATTGGTGGGTCAACTGGGGTTACCC	1057
Qy	937	ACCGAGTGGGGAAACAGAAAGCCCTGGTGAAGGTGGGACCCCTAGTGGTGACTTTGGC	996
Db	1058	ACCGAGTGGGGAAACAGAAAGCCCTGGTGAAGGTGGGACCCCTAGTGGTGACTTTGGC	1117
Qy	997	CGGGCTCAATGGCGGACTGGTTACGTCTCTCGGCCCCCTCCTCGAGAAATGGAGCC	1056
Db	1118	CGGGCTCAATGGCGGACTGGTTACGTCTCTCGGCCCCCTCCTCGAGAAATGGAGCC	1177
Qy	1057	AAGGTTGTCAGCTTCTTCAAGCAGCACGTCGCGGAGGTGAAGCACTGTACTCTGGGCTG	1116

Db	1178	AAAGTGTGACGCTTCTTCAAGCAGCAGCTGCCGGGAGGTGGAAAGCACTGTATCCTGGGCTG	1233
Qy	1117	GAGCGGCAACATTTCTCTTAAGAAAGTCCGAAAGAGAAATGACATGGCTCAAAATCTGCAA	1176
Db	1238	GAGGGGCACATTTCTCTTAAGAGTCCGAAAGGAGAAATGACATGGCTCAAAATCTGCAA	1297
Qy	1177	GGTGACCCGGCTGAGGATACCTCTTCTGCGCCCTGGCAAGAGCAGCTTAAAGCTTCCGAAA	1236
Db	1298	GGTGACCCGGCTGAGGATACCTCTTCTGCGCCCTGGCAAGAGCAGCTTAAAGCTTCCGAAA	1357
Qy	1237	GGCAATTCTCAAGAAAAGTCCCTACTCTGCTCAGGGGAGGTACAGGAGGACCCCTCAGGAA	1296
Db	1358	GGCAATTCTCAAGAAAAGTCCCTACTCTGCTCAGGGGAGGTACAGGAGGACCCCTCAGGAA	1417
Qy	1297	CTCAGACCGGTGCTCGATACTCCAGGGCAGCCTGCTCCCTGCTATCCCTGCTCCCAAGG	1356
Db	1418	CTCAGACCGGTGCTCGATACTCCAGGGCAGCCTGCTCCCTGCTATCCCTGCTCCCAAGG	1477
Qy	1357	AAAGGCATCCTTTAAGAAAGTCTCGACAGCTGAATCTGGTTACTACTCTCTCCAGAGCCC	1416
Db	1478	AAAGGCATCCTTTAAGAAAGTCTCGACAGCTGAATCTGGTTACTACTCTCTCCAGAGCCC	1537
Qy	1417	AGCGAGTCTGGGAACTCTTAGACGCGCAGTGAATGTTGTGAGTGGGGACCCCGTGGAG	1476
Db	1538	AGCGAGTCTGGGAACTCTTAGACGCGCAGTGAATGTTGTGAGTGGGGACCCCGTGGAG	1597
Qy	1477	CAGAAGTCTCCACAGGCTTCAGGGCTCCTCCTCCACCGCAAGGGCATTCCTCAAACTCAAT	1536
Db	1598	CAGAAGTCTCCACAGGCTTCAGGGCTCCTCCTCCACCGCAAGGGCATTCCTCAAACTCAAT	1657
Qy	1537	GGCAAGTCTCCCGCACAGCTTAGAGAGGCACTACCCCTAGCAGCTTTGGCTCCCTCGAC	1596
Db	1658	GGCAAGTCTCCCGCACAGCTTAGAGAGGCACTACCCCTAGCAGCTTTGGCTCCCTCGAC	1717
Qy	1597	CAACTGGCTCCTCCCATCTGTGACCGCGGCCACAGCGCGCCCTCAGGGGCTGTGAGTGAG	1656
Db	1718	CAACTGGCTCCTCCCATCTGTGACCGCGGCCACAGCGCGCCCTCAGGGGCTGTGAGTGAG	1777
Qy	1657	GACAGCATCTGTCTCCGAGTCTTTTGAACCAATTTGACCTGAACTGCTTTTCCCGAA	1716
Db	1778	GACAGCATCTGTCTCCGAGTCTTTTGAACCAATTTGACCTGAACTGCTTTTCCCGAA	1837
Qy	1717	ACCCCACTGAGGGGCTGTGTCTGTGCGAACAACTGAGGGGGCTTGAGCAGCCTCCCTCA	1776
Db	1838	ACCCCACTGAGGGGCTGTGTCTGTGCGAACAACTGAGGGGGCTTGAGCAGCCTCCCTCA	1897
Qy	1777	GAAGGTCTGAAGCGATGGTGCAGGAATCCTTTGGGGGATAGCTCTTTTCTGCACAGAC	1836
Db	1898	GAAGGTCTGAAGCGATGGTGCAGGAATCCTTTGGGGGATAGCTCTTTTCTGCACAGAC	1957
Qy	1837	TGCCAAGAGGTGACTGCAGCCTCAGACAGCCCTAGGATCTGCTCAAACTCAGC	1893
Db	1958	TGCCAAGAGGTGACTGCAGCCTCAGACAGCCCTAGGATCTGCTCAAACTCAGC	2014

RESULT 4	AY399041	1896 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	AY399041				
DEFINITION	Mus musculus HCM0091 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.				
ACCESSION	AY399041				
VERSION	AY399041.1	GI:39755030			
KEYWORDS	GSS				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
	1 (bases 1 to 1896)				
REFERENCE	Clark, A.G., Glancowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shtinsky, J.J., Adams, M.D. and Cargill, M.				
AUTHORS	Inferring nonneutral evolution from human-chimp-mouse orthologous				
TITLE					

gene trios  
Science 302 (5652), 1960-1963 (2003)  
14671302  
2 (bases 1 to 1896)  
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
Todd, M.A., Tanenbaum, D.M., Clivello, D.R., Lu, F., Murphy, B.,  
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
Direct Submission  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
Location/Qualifiers  
source  
1..1896  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
<1..>1896  
/locus\_tag="HCM0091"  
ORIGIN  
Query Match 93.7%; Score 1773; DB 9; Length 1896;  
Best Local Similarity 93.7%; Pred. No. 0;  
Matches 1773; Conservative 0; Mismatches 120; Indels 0; Gaps 0;  
Qy 1 ATGAGTCCGCTGGGCTTACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCCGGCCCTGGCC 60  
Db |  
Qy 1 ATGAGTCCGCTGGGCTTACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCCGGCCCTGGCC 60  
Db |  
Qy 61 TCGAGAGCGCCCGCGCTGGGGACGGCTCATCAAGTCCGCTAAACCTCTGATGAAG 120  
Db |  
Qy 61 TCGAGAGCGCCCGCGCTGGGGACGGCTCATCAAGTCCGCTAAACCTCTGATGAAG 120  
Db |  
Qy 121 AAGCAGCGGTGAAGCGGACCATCAAAACACAACTCGGGACCGCTACGAGTTCCTG 180  
Db |  
Qy 121 AAGCAGCGGTGAAGCGGACCATCAAAACACAACTCGGGACCGCTACGAGTTCCTG 180  
Db |  
Qy 181 GAGACGCTGGGCAAGGCACTACGGGAAGTGAAGAGGACGAGAGCTCGGGCGT 240  
Db |  
Qy 181 GAGACGCTGGGCAAGGCACTACGGGAAGTGAAGAGGACGAGAGCTCGGGCGT 240  
Db |  
Qy 241 CTGCTGATTAATCAAGTCCATCAGAAAGACAAATCAAGATGAGCAGGATCTGCTGAC 300  
Db |  
Qy 241 CTGCTGATTAATCAAGTCCATCAGAAAGACAAATCAAGATGAGCAGGATCTGCTGAC 300  
Db |  
Qy 301 ATACGAGGAGATGAGATCATGCTTCTCACTCAACCCACCATCATTTGCCATCAT 360  
Db |  
Qy 301 ATACGAGGAGATGAGATCATGCTTCTCACTCAACCCACCATCATTTGCCATCAT 360  
Db |  
Qy 361 GAAGTGTTCAGAAATAGCAGCAAGATTTGATTTGATGATGATGATGATGATGATGAT 420  
Db |  
Qy 361 GAAGTGTTCAGAAATAGCAGCAAGATTTGATTTGATGATGATGATGATGATGATGAT 420  
Db |  
Qy 421 CTGATGATTAATCAAGTCCATCAGAAAGACAAATCAAGATGAGCAGGATTTCTTC 480  
Db |  
Qy 421 CTGATGATTAATCAAGTCCATCAGAAAGACAAATCAAGATGAGCAGGATTTCTTC 480  
Db |  
Qy 481 CGACAGATCGTGTTCGCTGCTACTGCTACTGCTACTGCTACTGCTACTGCTACTGCT 540  
Db |  
Qy 481 CGACAGATCGTGTTCGCTGCTACTGCTACTGCTACTGCTACTGCTACTGCTACTGCT 540  
Db |  
Qy 541 AAGCTGGAACATCTCTCTAGATGCAATGGAACATCAAGATTTGCTGCTTTGGCTC 600  
Db |  
Qy 541 AAGCTGGAACATCTCTCTAGATGCAATGGAACATCAAGATTTGCTGCTTTGGCTC 600  
Db |  
Qy 601 TCCAACTGTAACCAAGCAAGTCTCTCCAGACGTTCTGTGGGAGCCCTCTCTACGCC 660  
Db |  
Qy 601 NNN 660  
Db |  
Qy 661 TCCCTGAGATAGTCAACGGAAGCCCTATGTGGGCCAGAGGTGACAGCTGCTGTCTG 720  
Db |  
Qy 661 NNN 720  
Db |

Qy 721 GGCGTTCTCTGTATACCTCTGGTGCATGGCACCATGCCCTTTGACGGCGCAGATCATAAA 780  
Db |  
Qy 721 GGCGTTCTCTGTATACCTCTGGTGCATGGCACCATGCCCTTTGACGGCGCAGATCATAAA 780  
Db |  
Qy 781 ACCTGTGTGAAGCAAAATCAGTAAACGGGCTTTACCGTGAGCGCCCAAGCCGCTCCGATGCC 840  
Db |  
Qy 781 ACCTGTGTGAAGCAAAATCAGTAAACGGGCTTTACCGTGAGCGCCCAAGCCGCTCCGATGCC 840  
Db |  
Qy 841 TGTGGCTGTATCCGGTGGCTGTTAATGGTGAACCCCAACCCGTCGGGCGACACTGGAGGAT 900  
Db |  
Qy 841 TGTGGCTGTATCCGGTGGCTGTTAATGGTGAACCCCAACCCGTCGGGCGACACTGGAGGAT 900  
Db |  
Qy 901 GTAGCCAGTCAATTTGGTGGTCAACTGGGTTTACACCAACCGAGTTCGGGGAACAGGAAGCC 960  
Db |  
Qy 901 GTAGCCAGTCAATTTGGTGGTCAACTGGGTTTACACCAACCGAGTTCGGGGAACAGGAAGCC 960  
Db |  
Qy 961 CTGCGTGAGGTTGGGCAACCTAGTGTGATCTTTGGCGGGCTCCATTTGGCGGACTGGTTA 1020  
Db |  
Qy 961 CTGCGTGAGGTTGGGCAACCTAGTGTGATCTTTGGCGGGCTCCATTTGGCGGACTGGTTA 1020  
Db |  
Qy 1021 CGTGGCTCTCGCGCCCTCTCTGGAGAAATGGAGCCAAAGTGTGCAGCTTCTTCAAGCAG 1080  
Db |  
Qy 1021 CGTGGCTCTCGCGCCCTCTCTGGAGAAATGGAGCCAAAGTGTGCAGCTTCTTCAAGCAG 1080  
Db |  
Qy 1081 CAGTGCCTGGGAGGTGGAAGCACTGTACCTGGGCTGGAGCGGCAACATTTCTTTAAGAG 1140  
Db |  
Qy 1081 CAGTGCCTGGGAGGTGGAAGCACTGTACCTGGGCTGGAGCGGCAACATTTCTTTAAGAG 1140  
Db |  
Qy 1141 TCCGAAAGGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCTCT 1200  
Db |  
Qy 1141 TCCGAAAGGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCTCT 1200  
Db |  
Qy 1201 TCTCGCCCTGGCAAGAGCAGCTTTAAGCTTCCGAAAGGCAATTTCTCAAGAAAGTCTCTCT 1260  
Db |  
Qy 1201 TCTCGCCCTGGCAAGAGCAGCTTTAAGCTTCCGAAAGGCAATTTCTCAAGAAAGTCTCTCT 1260  
Db |  
Qy 1261 ACCTGTGAGGAGGTGAGGAGGACCTCAGGAACTCAGACCGGCTCGCTGATCTCCA 1320  
Db |  
Qy 1261 ACCTGTGAGGAGGTGAGGAGGACCTCAGGAACTCAGACCGGCTCGCTGATCTCCA 1320  
Db |  
Qy 1321 GGCGAGCTGTCTCTGTATCTCTCCAGGAAAGGCAATCTTAAAGAAAGTCTCGA 1380  
Db |  
Qy 1321 GGCGAGCTGTCTCTGTATCTCTCCAGGAAAGGCAATCTTAAAGAAAGTCTCGA 1380  
Db |  
Qy 1381 CAGCGTGAATCTGGTTACTACTCTCTCAGAGCCAGCAGTCTGGGGAACCTTTAGAC 1440  
Db |  
Qy 1381 CAGCGTGAATCTGGTTACTACTCTCTCAGAGCCAGCAGTCTGGGGAACCTTTAGAC 1440  
Db |  
Qy 1441 GCCAGTGTGTGTTGTGAGTGGGACCCCGTGGAGCAGAAAGTCTCCACAGGCTTCAGGG 1500  
Db |  
Qy 1441 GCCAGTGTGTGTTGTGAGTGGGACCCCGTGGAGCAGAAAGTCTCCACAGGCTTCAGGG 1500  
Db |  
Qy 1501 CTCTCTCTCCACCGCAAGGCAATTTCTCAAACTCAATGGCAAGTCTCCCGCACAGCCTTA 1560  
Db |  
Qy 1501 CTCTCTCTCCACCGCAAGGCAATTTCTCAAACTCAATGGCAAGTCTCCCGCACAGCCTTA 1560  
Db |  
Qy 1561 GAAGGCACTACCCCTAGCAGCTTTGGCTCCCTGGACCAACTGGCTCTCTCCCATCTCTGCA 1620  
Db |  
Qy 1561 GAAGGCACTACCCCTAGCAGCTTTGGCTCCCTGGACCAACTGGCTCTCTCCCATCTCTGCA 1620  
Db |  
Qy 1621 GCCCGCCAGCGCCCTCAGGGCTGTGAGTGAGGACAGCATCTGTCTCCGAGTCC 1680  
Db |  
Qy 1621 GCCCGCCAGCGCCCTCAGGGCTGTGAGTGAGGACAGCATCTGTCTCCGAGTCC 1680  
Db |  
Qy 1681 TTTGACCAATTTGACCTTGTGCTGAACGCTTTCCGAAACCCCACTGAGGGGCTGTGTGCT 1740  
Db |  
Qy 1681 TTTGACCAATTTGACCTTGTGCTGAACGCTTTCCGAAACCCCACTGAGGGGCTGTGTGCT 1740  
Db |  
Qy 1741 GTGGACAACTGAGGGGCTTGTGAGGAGCTCTCTCAGAAAGTCTGAAGCGATGGTGGCAG 1800  
Db |  
Qy 1741 GTGGACAACTGAGGGGCTTGTGAGGAGCTCTCTCAGAAAGTCTGAAGCGATGGTGGCAG 1800  
Db |

```
QY 1801 GAATCTCTGGGGATAGTCTCTTCTCTGACAGATGCGCAAGAGTGACTGACGCTAC 1860
Db 1801 GAATCTCTGGGGATAGTCTCTTCTCTGACAGATGCGCAAGAGTGACTGACGCTAC 1860

QY 1861 AGACAAGCCCTAGGAATCTGCTCAAAAGCTCAGC 1893
Db 1861 AGACAAGCCCTAGGAATCTGCTCAAAAGCTCAGC 1893

RESULT 5
LOCUS AY399039
DEFINITION Homo sapiens HM0091 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY399039
VERSION AY399039.1 GI:39755028
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1887)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1887)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Direct Submission
COMMENT Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..1887
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1887
/locus_tag="HCM0091"
gene
ORIGIN
Query Match 66.6%; Score 1261; DB 9; Length 1887;
Best Local Similarity 78.4%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 380; Indels 21; Gaps 2;

QY 47 CCTCCGCGCTGGCTCGAGAGCGCCGCGCTGGCGGACCGGCTCATCAAGTCGCTA 106
Db 35 CCATCTCTCGGCGCAGAGCTAGCCCGCGCTGGCGGAGGCTGATCAAGTCGCCA 94

QY 107 AACTCTGATGAAGACGCGGTGAAGCGGACCATCAAAACAACTGGGCGACC 166
Db 95 AGCCCTTAATGAAGACGCGGTGAAGCGGACCATCAAAACAACTGGGCGACC 154

QY 167 GTACAGAGTTCTTGAGACGCTGGGCAAGGSCACTACGGGAAGGTGAAGAGCGACG 226
Db 155 GCTACAGAGTTCTTGAGACGCTGGGCAAGGSCACTACGGGAAGGTGAAGAGCGCGG 214

QY 227 AGAGCTCGGCGCTCTGGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGC 286
Db 215 AGAGCTCGGCGCTCTGGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGC 274

QY 287 AGGATCTGCTGCATACGAGGAGGATGAGATGCTTCACTCAACACACCCACCA 346
Db 275 AAGATCTGATGCATACGAGGAGGAGATGAGATGCTTCACTCAACACACCCCTACA 334
```

```
QY 347 TCATTGCCATCCATGAAGTGTGTTGAGAAATAGCAGCAAGATTTGTATGTCATGGAGTATG 406
Db 335 TCATTGCCATCCATGAAGTGTGTTGAGAAACAGCAGCAAGATCGTGATCGTCATGGAGTATG 394

QY 407 CCAGCGGAGCGGCGATCTGTATGATTAATCATCAGTGAAGCGGCACGGCTGAGTGAGCGGGAG 466
Db 395 CCAGCGGAGCGGCGACCTTTATGACTACATCAGCGAGCGGCGAGAGCTCACTGAGCGCGGAAG 454

QY 467 CCAGGCGATTTCTCCGACAGATCGTGTGCGCTGCACCTACTGCCACACAGAGCGGATCG 526
Db 455 CTAGGCGATTTCTCCGCGAGATCGTGTGCGCTGCACCTATTGGCCATCAGAACAGAGTTG 514

QY 527 TTCACCGAGATCTCAAGCTGGAAGAAATCCTTCTAGATGCCAATGGAACATCAAGATTTG 586
Db 515 TCCACCGAGATCTCAAGCTGGAAGAAATCCTTCTGATGCCAATGGAATATCAAGNNNN 574

QY 587 CTGACTTTGGCGCTCTCAACCTGTGACCAAAAGGCAAGTTCTCTCAGACGTTCTGTGGGA 646
Db 575 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 634

QY 647 GCCTCTCTACGCCTCGCCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCAGAGGTGG 706
Db 635 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 694

QY 707 ACAGCTGGTCTCTGGCGGTTCTCTGTATCATCTCTGTGATGCGACCATGCCCCCTTTGACG 766
Db 695 ACAGCTGGTCTCTGGCGGTTCTCTGTATCATCTCTGTGATGCGACCATGCCCCCTTTGATG 754

QY 767 GGAGGATCATAAACACCTGGTGAAGCAAAATCAGTAACGGGGTTACCGTGAGCGGCCCA 826
Db 755 GGAGGATCATAAACACCTGGTGAAGCAAAATCAGTAACGGGGTTACCGTGAGCGGCCCA 814

QY 827 AGCCGTCGCATGCTGTGCGCTGATCCGCTGGCTGTAAATGGTGAACCCACCCCGCTCGGG 886
Db 815 AACCTCTGATGCTGTGCGCTGATCCGCTGGCTGTAAATGGTGAACCCACCCCGCTCGGG 874

QY 887 CCACACTGGAGGATGTAGCAGTCAATTTGGTGGTCAACTGGGGTTACAACACCGAGTGC 946
Db 875 CCACCTGGAGGATGTAGCAGTCAATTTGGTGGTCAACTGGGGTTACAACACCGAGTGC 934

QY 947 GGGACAGGAAGCCCTCGCTGAGGGTGGGCAACCTAGTGGTGAATTTGGCGCGCTCCA 1006
Db 935 GAGAGCAGGAGGCTCCGCGATGAGGGTGGGCAACCTAGTGGTGAATTTGGCGCGCTCCA 994

QY 1007 TGGCGGACTGGTTACGTCGCTCTCGCGCCCTCTCTGGAGAAATGAGCAACAGGTGTGCA 1066
Db 995 TGGCTGACTGGTCCGCGCTCTCTCGCGCCCTCTCTGGAGAAATGAGCAACAGGTGTGCA 1054

QY 1067 GCTTCTTCAAGCAGCAGTCCGCGAGGTGAAGCACTGTACTCTGGGCTGGAGCGGCAAC 1126
Db 1055 GCTTCTTCAAGCAGCAGTCCGCGAGGTGAAGCACTGTACTCTGGGCTGGAGCGGCAAC 1114

QY 1127 ATTCTCTTAAGAGTCCGGAAGGAATCAGATGGCTCAAAATCTCAGAGGTGACCCGG 1186
Db 1115 ATTCTCTTAAGAGTCCGGAAGGAATCAGATGGCTCAAAATCTCAGAGGTGACCCGG 1174

QY 1187 CTGAGGATACCTCTCTCGCCCTGGCAAGCAGACCTTAAGCTTCCGAAAGGCAATCTCA 1246
Db 1175 CTGATGACTGCGCATCGCCCTGGCAAGCAGACCTTAAGCTTCCGAAAGGCAATCTCA 1234

QY 1247 AGAAAAGTCTCTCTA CTTCTGTAGGGAGGTA CAGGAGGACCTCTAGGAATCTCAGACCGG 1306
Db 1235 AGAAAAGTCTCTCAGCTCTGCAAGAGGGGTACAGGAGGACCTCTCAGGAGCTCAGCCCAA 1294

QY 1307 TGCTGTACTCTCAGGCGCAGCTGCTCCCTGCTATCTCTCTCTCCCAAGGAAAGGCAATCC 1366
Db 1295 TCCTTCTGAGCGCCAGGCGAGCTGCCCC-----CCCTGCTCTCCCAAGGAGGCAATTC 1345

QY 1367 TTAAGAGTCTCGACAGCGTGAATCTGGTTTACTCTCTCTCCAGAGCCCGAGGAGTCTG 1426
Db 1346 TCAAGAGCCCGCAGCAGCGAGTCTGGCTTACTCTCTCTCCGAGCCCGAGGAAATCTG 1405

QY 1427 GGGAACTCTTAGAGCGCAGTGTGTGTGTGAGTGGGGACCCCGGTGGAGCAGAGTCTC 1486
```





Db 1055 GCTTCTTCAAGCAGCATGACCTGGTGGGGGAAGCACCACCCCTGGCGCTGGAGCGCCAGC 1114

Qy 1127 ATTCTCTTAAGAGTCCCGAAGAGGAGNATCATGGCTCAAAATCTGCAAGGTGACCCGG 1186

Db 1115 ATTCTCTTAAGAGTCCCGAAGAGGAGNATCATGGCTCAAAATCTGCAAGGTGACCCGG 1174

Qy 1187 CTGAGGATACCTCTTCTCGCCCTGGCAAGAGCAGCTTAAAGCTTCGAAAGGATTTCTCA 1246

Db 1175 CTGATGACATGCCATCGCCCTGGCAAGAGCAGCTTAAAGCTTCGAAAGGATTTCTCA 1234

Qy 1247 AGAAAGAGTCTTACCTCGTTCAGGGAGGATGACAGGAGGACCTTCAGGAATCAGACCGG 1306

Db 1235 AGAAGAGGTGTGAGCTCTGCAAGAGGGGACACAGAGGAGCCTTCGAGCTCAGCCCAA 1294

Qy 1307 TGGCTGATATCAGGAGCAGCTGCTCCTGCTGATCTCCTGCTTCCCAAGGAAGGATTC 1366

Db 1295 TCCCTGAGAGCCAGGAGGAGCTGCC-----CCCTGCTCCCCAAGAGGAGGATTC 1345

Qy 1367 TTAAGAGTCTGACAGCAGTGAATCTGTTACTACTCTCTCCAGAGCCAGCGAGTCTG 1426

Db 1346 TCAGAGAGCCCAACAGCGGAGTCTGCTACTACTCTCTCCAGAGCCAGTGAATCTG 1405

Qy 1427 GGAAGTCTTACAGCGCAGTGTGTTGTGAGTGGGAGCCCGCTGGAGCAGAGTCTC 1486

Db 1406 GGGAGCTCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1465

Qy 1487 CAGAGCTTCAAGGCTCTCTCCAGCGAAGGAGCTTCAAACTCAATGGCAAGTCT 1546

Db 1466 NNN 1525

Qy 1547 CCGCAGCAGCTTTAGAGGAGCAGTACCCCTAGCAGCTTTGGCTCTCGTCCGACCAACTGGCCT 1606

Db 1526 NNNNAGCAGCTTTGAGCTCGGCGCCCGCAGCAGCTTTGGCTCTCGTCCGACCAACTGGCCT 1585

Qy 1607 CTTCCCACTCTGAGCGCCCGCAGCGCCCTCAGGGGCTGTGAGTGGAGCAGCATCC 1666

Db 1586 CATCTCGCCCTCGGCGCGGCGCAGCGAGCCCTCAGGGGCTGTGAGCGAGCAGCATCC 1645

Qy 1667 TGTCTCCGAGTCTTTGACCAATTTGACTTGTGCTGAGCAGTCTTCCGGAACCCCACTGA 1726

Db 1646 TGTCTCTGAGTCTTTGACCAATTTGACTTGTGCTGAGCAGTCTTCCGGAACCCCACTGC 1705

Qy 1727 GGGGCTGTGTGTGTGTGAGCAACCTGAGGGGCTTTGAGAGCCTTCCCTCAGAAAG----- 1780

Db 1706 GGGGCTGTGTGTGTGTGAGCAACCTGAGGGGCTTTGAGAGCCTTCCCTCAGAGGGCCCTG 1765

Qy 1781 -----GTCTGAAGCGATGTGGCAGGAATCTTTGGGGGATAGTGTCTTTCTGACAG 1834

Db 1766 GAAGCTGCTGAGGCGCTGGCGGAGGATCTTTGGGGGAGCAGTGTCTTTTCCCTGACAG 1825

Qy 1835 ACTGCCAAGAGTCACTGACGCTTACAGACAGCCCTAGGAATCTGCTCAAGCTCAGC 1893

Db 1826 ACTGCCAAGAGTCACTGACGCTTACAGACAGCCCTAGGAATCTGCTCAAGCTCAGC 1884

RESULT 7  
 CB248251  
 LOCUS CB248251 726 bp mRNA linear EST 15-JUL-2003  
 DEFINITION UI-M-FD0-byi-a-21-0-UI.r1 NIH BMAP\_F00 Mus musculus cDNA clone  
 IMAGE:5718428 5', mRNA sequence.  
 ACCESSION CB248251  
 VERSION CB248251.1 GI:28386432  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 726)  
 AUTHORS NTH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-x@mail.nih.gov](mailto:cgapbs-x@mail.nih.gov)

Tissue Procurement: Dr. James Lin, University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
 This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)

Seq primer: pYX-5.  
 Location/Qualifiers  
 1. 726  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5718428"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 12.5 dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="NIH\_BMAP\_F00"  
 /notes="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;  
 Site 2: Not I; The library was constructed according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured mRNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with an  
 oligo-dT primer containing a Not I site. Double stranded  
 cDNA was size selected according to mRNA size fraction,  
 ligated with EcoR I adaptor, digested with Not I, and then  
 cloned directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA  
 tail, is TGAGAGAGCC. This library was created for the  
 University of Iowa Mouse Brain Molecular Anatomy Project  
 (BMAP). Gene Discovery in the Developing Mouse Nervous  
 System", supported by National Institutes of Mental Health  
 (NIMH), Hemin Chin, Ph.D., program coordinator."

## ORIGIN

Query Match 36.1%; Score 683.6; DB 6; Length 726;  
 Best Local Similarity 98.8%; Pred. No. 2.8e-167;  
 Matches 720; Conservative 0; Mismatches 5; Indels 4; Gaps 3;  
 Qy 800 GTAAACGGGGCTTAC-CGTGAGCGCGCCCAAGCCCTCCGATCGCTGTGGCTGTGATCCGGTGG 858  
 Db 1 GTAAACGGGGCTTACNCGTGAAGCGCCCAAGCCCTCCGATCGCTGTGGCTGTGATCCGGTGG 60  
 Qy 859 CTGTTAATGTGTAACCCACCCCGTGGGGCCACACTGAGGATGATGAGCAGTCAATGGTGG 918  
 Db 61 CTGTTAATGTGTAACCCACCCCGTGGGGCCACACTGAGGATGATGAGCAGTCAATGGTGG 120  
 Qy 919 GTCAACTGGGGTTACACCCCGGAGTCGGGGAACAGGAAGCCCTCGGTGAGGGTGGGCAC 978  
 Db 121 GTCAACTGGGGTTACACCCCGGAGTCGGGGAACAGGAAGCCCTCGGTGAGGGTGGGCAC 180  
 Qy 979 CCTAGTGGTGAATTTGGCGGGGCTCCATGGCGGAGTGTGTTACGTGCTCCTCGCGCCC 1038  
 Db 181 CCTAGTGGTGAATTTGGCGGGGCTCCATGGCGGAGTGTGTTACGTGCTCCTCGCGCCC 240  
 Qy 1039 CTCTGGAGATCGAGCAAGGTGTGACAGTCTTCTCAAGCAGCAGCTGCCGGGAGTGA 1098  
 Db 241 CTCTGGAGATCGAGCAAGGTGTGACAGTCTTCTCAAGCAGCAGCTGCCGGGAGTGA 300  
 Qy 1099 AGCACTGTACTGGGCTGGAGCGGCAACATTTCTTAAAGATCCCGAAAGGAGATGAC 1158  
 Db 301 AGCACTGTACTGGGCTGGAGCGGCAACATTTCTTAAAGATCCCGAAAGGAGATGAC 360  
 Qy 1159 ATGGCTCAAAATCTCAAGAGTGACCGGTGAGGATACCTCTTCTGCGCTCGCAAGAGC 1218  
 Db 361 ATGGCTCAAAATCTCAAGAGTGACCGGTGAGGATACCTCTTCTGCGCTCGCAAGAGC 420  
 Qy 1219 AGCCTTAAGTTCGAAAGGCATTTCTCAAGAAAAGTCTCTACCTCGTCAGGGGAGGTA 1278  
 Db 421 AGCCTTAAGTTCGAAAGGCATTTCTCAAGAAAAGTCTCTACCTCGTCAGGGGAGGTA 480



QY 1279 CAGGAGACCTCAGGAACCTCAGACCGGTGCTGATCTCCAGGCGAGCCTGTCCCTGCT 1338  
 |||||  
 Db 481 CAGGAGACCTCAGGAACCTCAGACCGGTGCTGATCTCCAGGCGAGCCTGTCCCTGCT 540  
 |||||  
 QY 1339 GTATCCCTGCTCCCAAGAAAGGATCTTAAAGATCTCCAGACGCTGAATCTGGTTAC 1398  
 |||||  
 Db 541 GTATCCCTGCTCCCAAGAAAGGATCTTAAAGATCTCCAGACGCTGAATCTGGTTAC 600  
 |||||  
 QY 1399 TACTCCCTCTCCAGAGCCAGGAGTCTGGGAACTCTTAGACGCCAGTGATGTTGTG 1458  
 |||||  
 Db 601 TACTCCCTCTCCAGAGCCAGGAGTCTGGGAACTCTTAGACGCCAGTGATGTTGTG 660  
 |||||  
 QY 1459 AGTGGGACCCCGTGGAGCAGAACTCTCCACAGGCTTCAGGCTCCCTCCACCGCAAG 1518  
 |||||  
 Db 661 AGCGGGA-CCCGTGGAGCAGAA--TCTCACAGGCTTCAGGCTCCCTCCACCGCAAG 717  
 |||||  
 QY 1519 GGCATTCTC 1527  
 |||||  
 Db 718 GGCATTCTC 726  
 |||||

## RESULT 8

CO424322  
 LOCUS  
 DEFINITION UI-M-HUO-cqx-a-14-0-UI.r1 NIH\_BMAP\_HUO Mus musculus cDNA clone  
 IMAGE:30665389 5', mRNA sequence.

ACCESSION CO424322

VERSION CO424322.1 GI:49670481

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

## ORGANISM

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 672)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Dr. James Lin University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/mousefl.html>

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..672  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:30665389"  
 /tissue\_type="whole eye"  
 /dev\_stage="newborn (1, 5, 15 days)"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="NIH\_BMAP\_HUO"  
 /note="Organ: Eye; Vector: pYX-Asc; Site 1: Ecor I;  
 Site 2: Not I; The library was constructed according  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated  
 with Ecor I adaptor, digested with NotI and then cloned  
 directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is AATTAATACG. This library was created for the University  
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
 Developing Mouse Nervous System', supported by National

## FEATURES

source

Institute of Mental Health (NIMH)."

Query Match 35.5%; Score 672; DB 7; Length 672;  
 Best Local Similarity 100.0%; Pred. No. 3e-164;  
 Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 791 AGCAAAATCAGTAACGGGGCTTACCGTCAGCGCGCCCAAGCGTCCGATGCCGTGTGGCCCTGA 850  
 |||||  
 Db 1 AGCAAAATCAGTAACGGGGCTTACCGTCAGCGCGCCCAAGCGTCCGATGCCGTGTGGCCCTGA 60  
 |||||  
 QY 851 TCCGGTGGCTGTTAATGGTGAACCCACCCCGTGGGCCACACATGGAGGATGTAGCCAGTC 910  
 |||||  
 Db 61 TCCGGTGGCTGTTAATGGTGAACCCACCCCGTGGGCCACACATGGAGGATGTAGCCAGTC 120  
 |||||  
 QY 911 ATTGGTGGGTCAACTGGGGTTACACACCGGAGTCGGGGACAGGAAGCCCTGGTGGAGG 970  
 |||||  
 Db 121 ATTGGTGGGTCAACTGGGGTTACACACCGGAGTCGGGGACAGGAAGCCCTGGTGGAGG 180  
 |||||  
 QY 971 GTGGGCACCCCTAGTGGTGACTTTGGCCGGGCTCCATGGCGGACTGGTTAGTCGCTCCT 1030  
 |||||  
 Db 181 GTGGGCACCCCTAGTGGTGACTTTGGCCGGGCTCCATGGCGGACTGGTTAGTCGCTCCT 240  
 |||||  
 QY 1031 CGGCCCCCTCTCTGGAGAATGGAGCCAAAGTGTGCACTTCTTCAAGCAGCACATGTCGGG 1090  
 |||||  
 Db 241 CGGCCCCCTCTCTGGAGAATGGAGCCAAAGTGTGCACTTCTTCAAGCAGCACATGTCGGG 300  
 |||||  
 QY 1091 GAGGTGGAAGCACTGTACTTGGCTGGAGCGGCACACATTCTCTTAAGAAGTCCCGAAGG 1150  
 |||||  
 Db 301 GAGGTGGAAGCACTGTACTTGGCTGGAGCGGCACACATTCTCTTAAGAAGTCCCGAAGG 360  
 |||||  
 QY 1151 AGAATGACATGGCTCAAAATCTGCAAGTGACCCGGCTGAGGATACCTCTTCTCGCCCTG 1210  
 |||||  
 Db 361 AGAATGACATGGCTCAAAATCTGCAAGTGACCCGGCTGAGGATACCTCTTCTCGCCCTG 420  
 |||||  
 QY 1211 GCAAGAGCAGCCTTAAGCTTCCGAAAGGCATTTCTCAAGAAAAAGTCTCTTACCTCGTCAG 1270  
 |||||  
 Db 421 GCAAGAGCAGCCTTAAGCTTCCGAAAGGCATTTCTCAAGAAAAAGTCTCTTACCTCGTCAG 480  
 |||||  
 QY 1271 GGGAGGTACAGGAGGACCTTCAGGACTCAGCCGGTGCCTGATCTCCAGGGCAGCCTG 1330  
 |||||  
 Db 481 GGGAGGTACAGGAGGACCTTCAGGACTCAGCCGGTGCCTGATCTCCAGGGCAGCCTG 540  
 |||||  
 QY 1331 TCCCTGCTGTATCCCTCTCCCAAGGAAAGGCATCTTAAAGAGTCTCGACAGCGTGAAT 1390  
 |||||  
 Db 541 TCCCTGCTGTATCCCTCTCCCAAGGAAAGGCATCTTAAAGAGTCTCGACAGCGTGAAT 600  
 |||||  
 QY 1391 CTGGTTACTACTCTCTCCAGAGCCAGCGAGTCGGGAACTCTTAGAGCCCAAGTCATG 1450  
 |||||  
 Db 601 CTGGTTACTACTCTCTCCAGAGCCAGCGAGTCGGGAACTCTTAGAGCCCAAGTCATG 660  
 |||||  
 QY 1451 TGTTTGTGAGTG 1462  
 |||||  
 Db 661 TGTTTGTGAGTG 672  
 |||||

## RESULT 9

CK781308

LOCUS

DEFINITION

UI-M-GIO-clp-g-23-0-UI.r1 NIH\_BMAP\_G10 Mus musculus cDNA clone

IMAGE:30619078 5', mRNA sequence.

CK781308

VERSION

CK781308.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 655)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov  
Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/mouse1.html  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. .655  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30619078"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH BMAP G10"

/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

# FEATURES

source

## RESULT 10

BG174288

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BG174288 805 bp mRNA linear EST 06-FEB-2001  
602334589F1 NCI\_COAP\_Mam1 Mus musculus cDNA clone IMAGE:4457398 5',  
mRNA sequence.

BG174288.1 GI:13680991

EST.

Mus musculus (house mouse)

Mus musculus

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NIH-MGC http://mgs.nci.nih.gov/

1 (bases 1 to 805)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLNL0254 row: b column: 23

High quality sequence stop: 704.

Location/Qualifiers

1. .805

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:4457398"

/tissue\_type="tumor, biopsy sample"

/dev\_stage="3 months, virgin"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Mam1"

/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI. Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

## FEATURES

source

## ORIGIN

Query Match 34.4%; Score 651.6; DB 4; Length 805;  
Best Local Similarity 96.6%; Pred. No. 6.9e-159;  
Matches 709; Conservative 0; Mismatches 19; Indels 6; Gaps 4;  
QY 379 AGCAAGATTGTGTTGTCATGGAGTATGCCAGCGGCGGATCTGTGATTACATCAGT 438  
DB 1 AGCAAGATTGTGTTGTCATGGAGTATGCCAGCGGCGGATCTGTGATTACATCAGT 60  
QY 439 GAGCGGCCAGCGGTGAGTGAGCGGAGCCGAGGATTTCTCCAGACAGATCGTCTCGCC 498  
DB 61 GAGCGGCCAGCGGTGAGTGAGCGGAGCCGAGGATTTCTCCAGACAGATCGTCTCGCC 120  
QY 499 CTGCACTACTGCCACCAAGAGCGGATCGTTTCCCGAGATCTCAAGCTTGGAACATCCTT 558  
DB 121 CTGCACTACTGCCACCAAGAGCGGATCGTTTCCCGAGATCTCAAGCTTGGAACATCCTT 180

```
Qy 559 CTAGATGCCAATGGAACATCAAGATTGCTGACTTTGGCCCTCTCAACCTGTACCAAAA 618
Db 181 CTAGATGCCAATGGAACATCAAGATTGCTGACTTTGGCCCTCTCAACCTGTACCAAAA 240
Qy 619 GGCAGTTCTCCAGACGTTCTGTGGAGCCCTCTACGGCTCGCTGAGATAGTCAAC 678
Db 241 GGCAGTTCTCCAGACGTTCTGTGGAGCCCTCTCTACGGCTCGCTGAGATAGTCAAC 300
Qy 679 GGAAGCCCTATGTGGGCGCAGAGGTGGACAGTGGTCTCTGGCGTTCTCTCTGATC 738
Db 301 GGAAGCCCTATGTGGGCGCAGAGGTGGACAGTGGTCTCTGGCGTTCTCTCTGATC 360
Qy 739 CTGTGTGATGCACCATGCTCTTTGACGGGAGGATCATAAACACTGGTGAACAAATC 798
Db 361 CTGTGTGATGCACCATGCTCTTTGACGGGAGGATCATAAACACTGGTGAACAAATC 420
Qy 799 AGTAACGGGGTTTACCTGAGCCGCCAGCCGCTCGATGCTGTGGCTGATCCGGTGG 858
Db 421 AGTAACGGGGTTTACCTGAGCCGCCAGCCGCTCGATGCTGTGGCTGATCCGGTGG 480
Qy 859 CTGTTAATGTGTGAACCCACCCCGTGGGCCACACTGGAGGATGTAGCCAGTCAATTGGTGG 918
Db 481 CTGTTAATGTGTGAACCCACCCCGTGGGCCACACTGGAGGATGTAGCCAGTCAATTGGTGG 540
Qy 919 GTCAACTGGGTTTACACACCGGAGTGGGGAAACAGGAAGCCCTGCGTGAGGGTGGGCAC 978
Db 541 GTCAACTGGGTTTACACACCGGAGTGGGGAAACAGGAAGCCCTGCGTGAGGGTGGGCAC 600
Qy 979 CTTAGTGTGACTTTGGCGGGCTCTCATGCGGACTGGTTACGTGCTCTCGCGCC 1038
Db 601 CTTAGTGTGACTTTGGC--GGGCTCATGCGGACTGGTTACGTGCTC--TCGCGCC 656
Qy 1039 CTCCTGAGAAATGAGCAAGCTGTGACAGTCTTCAAGCAGACGTCGCGGAGGTGA 1098
Db 657 CTCCTGAGAAATGG--GCCAGTGTGACAGTCTTCAAGCAGACGTCGCGGAGGTGG 714
Qy 1099 AGCACTGTACTGG 1112
Db 715 AGCACTGTACTGG 728
```

```
RESULT 11
BG694881
LOCUS
DEFINITION NISC iv09h10.w1 Soares NMBP2 pituitary Mus musculus cDNA clone
IMAGE: 4317546 5', mRNA sequence.
ACCESSION BG694881
VERSION BG694881.1 GI:13954237
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 646)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
cDNA Library Preparation: M. Bento Soares Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
MGI:1598314
Plate: LLAM9921 row: O column: 19
Seq primer: T7 primer.
Location/Qualifiers
1. .646
```

FEATURES  
source

```
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:4317546"
/tissue_type="pituitary gland"
/dev_stage="embryo, 14 dpc"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares NMBP2 pituitary"
/notes="Organ: brain; Vector: pT73D-Pac; Site 1: NotI;
Site 2: EcoRI; 1st strand cDNA was primed with a NotI -
oligo(dT) primer
5'-AATCGGAAGAAATTCGCGCGCGCGCTTTTCTTTTCTTTTCTTTT-3';
double-stranded cDNA was ligated to EcoRI adaptors
5'-AATTCGGCAGCAGG-3' AND 5'-CCTCGTGGC-3' (Pharmacia),
digested with NotI and cloned into the NotI and EcoRI
sites of the pT73D-PacI vector. Library went through one
round of normalization, and was constructed in the
laboratory of M. Bento Soares (University of Iowa)."
```

## ORIGIN

```
Query Match 34.1% Score 646; DB 4; Length 646;
Best Local Similarity 100.0%; Pred. No. 1.9e-157; Indels 0; Gaps 0;
Matches 646; Conservative 0; Mismatches 0;
Qy 1193 ATACCTCTTCTCGCCCTGGCAAGAGCAGCCCTTAAGCTTCCGAAGGCAATCTCAAGAAAA 1252
Db 1 ATACCTCTTCTCGCCCTGGCAAGAGCAGCCCTTAAGCTTCCGAAGGCAATCTCAAGAAAA 60
Qy 1253 AGTCCTCTACCTGTGAGGGAGGTACAGAGAGACCTCAGGAACCTCAGACCGGTGCGTG 1312
Db 61 AGTCCTCTACCTGTGAGGGAGGTACAGAGAGACCTCAGGAACCTCAGACCGGTGCGTG 120
Qy 1313 ATACTCCAGGGCAGCTGCTCCCTGTATCTCTGCTATCTCCCAAGAAAGGCAATCTTAAGA 1372
Db 121 ATACTCCAGGGCAGCTGCTCCCTGTATCTCTGCTATCTCCCAAGAAAGGCAATCTTAAGA 180
Qy 1373 AGTCTCGACAGCGTGAATCTGGTTACTACTCTCTCCAGAGCCCGAGAGTCTGGGGAAC 1432
Db 181 AGTCTCGACAGCGTGAATCTGGTTACTACTCTCTCCAGAGCCCGAGAGTCTGGGGAAC 240
Qy 1433 TCTTAGACCGCAGTGATGTGTTGTGAGTGGGACCCCGTGGAGCAGAGTCTCCACAGG 1492
Db 241 TCTTAGACCGCAGTGATGTGTTGTGAGTGGGACCCCGTGGAGCAGAGTCTCCACAGG 300
Qy 1493 CTTTCAGGGCTCTCTCCACCGCAAGGGCAATCTCAAACTCAATGCAAGTCTTCCCGCA 1552
Db 301 CTTTCAGGGCTCTCTCTCCACCGCAAGGGCAATCTCAAACTCAATGCAAGTCTTCCCGCA 360
Qy 1553 CAGCCTTAGAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCCTCTCTCC 1612
Db 361 CAGCCTTAGAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCCTCTCTCC 420
Qy 1613 ATCTGAGCCCGCCGCGCCCTCAGGGGCTGTGAGTGGAGCAGCATCTGTCTCT 1672
Db 421 ATCTGAGCCCGCCGCGCCCTCAGGGGCTGTGAGTGGAGCAGCATCTGTCTCT 480
Qy 1673 CCAGTCTCTTTGACCAATTTGGACTTGCCTGAAAGTCTTCCCGAAACCCCACTGAGGGGCT 1732
Db 481 CCAGTCTCTTTGACCAATTTGGACTTGCCTGAAAGTCTTCCCGAAACCCCACTGAGGGGCT 540
Qy 1733 GTGTGTCTGTGGACAACTCAGGGGGCTTGGAGCAGCTCCCTCAGAGGCTCTGAAGCGAT 1792
Db 541 GTGTGTCTGTGGACAACTCAGGGGGCTTGGAGCAGCTCCCTCAGAGGCTCTGAAGCGAT 600
Qy 1793 GTTGGCAGGAATCTTGGGGGATAGCTGCTTTTCTCTGACAGACTG 1838
Db 601 GTTGGCAGGAATCTTGGGGGATAGCTGCTTTTCTCTGACAGACTG 646
```

## RESULT 12

```
BF789245
LOCUS
DEFINITION 602105144F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4223267
```

5', mRNA sequence.  
BF789245  
BF789245.1 GI:12094281  
EST.  
Mus musculus (house mouse)  
Mus musculus  
ORGANISM  
REFERENCE  
1 (bases 1 to 948)  
NTH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs@remail.nih.gov  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLAM9811 row: g column: 12  
High quality sequence stop: 651.  
Location/Qualifiers  
1..948  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4223267"  
/lab\_host="NCI CGAP Kid4"  
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library. |"  
FEATURES  
source  
Query Match 34.1%; Score 645.8; DB 2; Length 948;  
Best Local Similarity 96.3%; Pred. No. 2.3e-157;  
Matches 683; Conservative 0; Mismatches 22; Indels 4; Gaps 2;  
QY 867 GGTGAACCCACCGCTCGGGCCACACTCGAGGATGTAGCCAGTCAATGGTGGTCAACTG 926  
DB 12 GGTGAACCCACCGCTCGGGCCACACTCGAGGATGTAGCCAGTCAATGGTGGTCAACTG 71  
QY 927 GGGTTACACCCGAGTCCGGGAAACAGGAAGCCCTGCTGAGGTGGGCAACCTAGTGG 986  
DB 72 GGGTTACACCCGAGTCCGGGAAACAGGAAGCCCTGCTGAGGTGGGCAACCTAGTGG 131  
QY 987 TGACTTTGGCGGGCTCATGGCGGACTGGTACGTCTGCTCGCGCCCTCCCTGGA 1046  
DB 132 TGACTTTGGCGGGCTCATGGCGGACTGGTACGTCTGCTCGCGCCCTCCCTGGA 191  
QY 1047 GAATGGAGCAAGGTGTGAGCTTCTTCAAGCAGCAGCTGCGGGAGGTGGAAGCACTGT 1106  
DB 192 GAATGGAGCAAGGTGTGAGCTTCTTCAAGCAGCAGCTGCGGGAGGTGGAAGCACTGT 251  
QY 1107 ACCTGGGTGGAGCGGCAACATCTCTTAAAGATCTCCGAAAGGAGATGATGCTGCTCA 1166  
DB 252 ACCTGGGTGGAGCGGCAACATCTCTTAAAGATCTCCGAAAGGAGATGATGCTGCTCA 311  
QY 1167 AAATCTGAAGGTGACCGGCTGAGGATACCTTCTCGCCCTGCGAGAGCAGCTTAA 1226  
DB 312 AAATCTGAAGGTGACCGGCTGAGGATACCTTCTCGCCCTGCGAGAGCAGCTTAA 371  
QY 1227 GCTTCCGAAAGGCAATCTTCAAGAAAAGTCTCTACCTCGTCAGGGAGGTACAGGAGA 1286  
DB 372 GCTTCCGAAAGGCAATCTTCAAGAAAAGTCTCTACCTCGTCAGGGAGGTACAGGAGA 431  
QY 1287 CCCTCAGGAACTCAGACCGGTGCTGATCTCAGGGAGCCTGTCCTGCTGTATCCCT 1346  
DB 432 CCCTCAGGAACTCAGACCGGTGCTGATCTCAGGGAGCCTGTCCTGCTGTATCCCT 491

QY 1347 GCTCCCAAGGAAGGCATCTTAAGACTCTCGACAGCGTGAATCTGTTACTCTCTC 1406  
DB 492 GCTCCCAAGGAAGGCATCTTAAGACTCTCGACAGCGTGAATCTGTTACTCTCTC 551  
QY 1407 TCCAGAGCCAGCGAGTCTGGGGAACTCTTAGACCCAGTGAATGTGTTGTAGTGGGA 1466  
DB 552 TCCAGAGCCAGCGAGTCTGGGGAACTCTTAGACCCAGTGAATGTGTTGTAGTGGGG 611  
QY 1467 CCGGTGGAGCAGAAGTCTTCCACAGCTTCAGGGTCTCTCCACCGCAAGGGCATCT 1526  
DB 612 CCGGTGGAGCAGAAGTCTTCCACAGCTTCAGGGTCTCTCCACCGCAAGGGCATCT 669  
QY 1527 GAACTCAATGCAAGTCTCTCCGACAGCTTGAAGGCACTACCCCT 1575  
DB 670 CAA--CTCATGCAAGTCTCTCCGACAGCTTAAAGGCTACCTAGCACT 716  
RESULT 13  
LOCUS  
DEFINITION  
CA319312 UI-M-FW0-cbz-i-16-0-UI.r1 NIH\_BMAP\_FW0 Mus musculus cDNA clone  
IMAGE:6816569 5', mRNA sequence.  
ACCESSION  
VERSION  
CA319312.1 GI:24537436  
KEYWORDS  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
REFERENCE  
1 (bases 1 to 795)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS  
NTH-MGC http://mgc.nci.nih.gov/  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished (1999)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs@remail.nih.gov  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)  
Seq primer: pYX-5.  
Location/Qualifiers  
1..795  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6816569"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH\_BMAP\_FW0"  
/notes="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGCAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 20:17:54 ; Search time 989.334 Seconds  
(without alignments)  
10994.228 Million cell updates/sec

Title: US-09-980-464-4\_COPY\_123\_2015  
Perfect score: 1893  
Sequence: 1 atggatcggtgcttact.....gaatctgtcaagctcagc 1893

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues  
Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA: \*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq: \*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq: \*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq: \*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq: \*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq: \*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq: \*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq: \*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq: \*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq: \*  
18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq: \*  
19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq: \*  
20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq: \*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1893	100.0	2902	15	US-10-355-975-4
2	1893	100.0	3073	17	US-10-322-281-22
3	1888.2	99.7	2027	17	US-10-343-514-87
4	1883.4	88.9	2026	17	US-10-343-514-27
5	1683.4	88.9	2929	17	US-10-343-514-1
6	1360.2	71.9	1884	9	US-09-963-159-3
7	1360.2	71.9	1884	16	US-10-423-543-45
8	1360.2	71.9	3353	9	US-09-963-159-1
9	1360.2	71.9	3353	16	US-10-423-543-43
10	1357.6	71.7	3360	16	US-10-311-034-38
11	1357	71.7	3404	17	US-10-322-281-25
12	1357	71.7	3463	18	US-10-618-941-11

13	1355.4	71.6	2501	16	US-10-302-172-215	Sequence 215, App
14	1355.4	71.6	3443	18	US-10-370-715B-639	Sequence 639, App
15	1071.4	56.6	1186	17	US-10-343-514-13	Sequence 13, Appl
16	1059.4	56.0	37278	17	US-10-322-281-21	Sequence 21, Appl
17	943.4	49.8	1186	17	US-10-343-514-70	Sequence 70, Appl
18	840.2	44.4	2616	17	US-10-363-829-151	Sequence 151, App
19	840.2	44.4	2619	17	US-10-363-829-250	Sequence 250, App
20	723	38.2	39699	17	US-10-322-281-24	Sequence 24, Appl
21	648	34.2	707	17	US-10-343-514-91	Sequence 91, Appl
22	648	34.2	712	17	US-10-343-514-31	Sequence 31, Appl
23	644.2	34.0	661	17	US-10-343-514-93	Sequence 93, Appl
24	613.6	32.4	741	17	US-10-343-514-32	Sequence 32, Appl
25	592	31.3	734	17	US-10-343-514-92	Sequence 92, Appl
26	583.2	30.8	588	17	US-10-343-514-97	Sequence 97, Appl
27	557.4	29.4	660	17	US-10-343-514-33	Sequence 33, Appl
28	533	28.2	578	17	US-10-343-514-95	Sequence 95, Appl
29	515	27.2	548	17	US-10-343-514-96	Sequence 96, Appl
30	510.6	27.0	579	17	US-10-343-514-35	Sequence 35, Appl
31	506.4	26.8	1986	9	US-09-780-949-3	Sequence 5, Appli
32	506.4	26.8	2884	9	US-09-780-949-5	Sequence 1, Appli
33	506.4	26.8	6828	9	US-09-780-949-1	Sequence 81, Appl
34	506.4	26.8	6828	15	US-10-354-358-81	Sequence 1541, Ap
35	506.4	26.8	6828	15	US-10-172-118-1541	Sequence 1541, Ap
36	506.4	26.8	6828	16	US-10-342-887-1541	Sequence 2552, Ap
37	506.4	26.8	6828	18	US-10-723-860-2552	Sequence 6703, Ap
38	506.4	26.8	6854	18	US-10-723-860-6703	Sequence 37, Appl
39	493.6	26.1	585	17	US-10-343-514-37	Sequence 36, Appl
40	483.8	25.6	548	17	US-10-343-514-36	Sequence 94, Appl
41	448.2	23.7	521	17	US-10-343-514-94	Sequence 34, Appl
42	417.8	22.1	521	17	US-10-343-514-34	Sequence 10, Appl
43	383.4	20.3	439	17	US-10-343-514-10	Sequence 14, Appl
44	381.4	20.1	436	17	US-10-343-514-14	Sequence 9, Appli
45	365.6	19.3	421	17	US-10-343-514-9	

ALIGNMENTS

RESULT 1  
US-10-355-975-4  
; Sequence 4, Application US/10355975  
; Publication No. US20030162277A1  
; GENERAL INFORMATION:  
; APPLICANT: Immunex Corporation  
; APPLICANT: Bird, Timothy A.  
; APPLICANT: Virca, G. Duke  
; APPLICANT: Martin, Unja  
; APPLICANT: Anderson, Dirk M.  
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES  
; FILE REFERENCE: 2923-A  
; CURRENT APPLICATION NUMBER: US/10/355,975  
; CURRENT FILING DATE: 2003-01-30  
; PRIOR APPLICATION NUMBER: US/09/579,664B  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 2902  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-355-975-4

Query Match	100.0%;	Score 1893;	DB 15;	Length 2902;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1893;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	ATGGAGTGGTGGCTTACTCCAGCGCCGAGCAGGCTCCCTCGGCTCGGCCCTGGCC	60	
Db	123	ATGGAGTGGTGGCTTACTCCAGCGCCGAGCAGGCTCCCTCGGCTCGGCCCTGGCC	182	
Qy	61	TCGGAGAGCGCCGGCGCTGGCGGACGGCTCATCAAGTCGGCTAAACCTCTGATGAAG	120	
Db	183	TCGGAGAGCGCCGGCGCTGGCGGACGGCTCATCAAGTCGGCTAAACCTCTGATGAAG	242	



121 AAGCAGCGGTGAGCGGCAACCATCAAAACACAACTCGGGCACCGCTACGATTCCTG 180  
Db  
243 AAGCAGCGGTGAGCGGCAACCATCAAAACACAACTCGGGCACCGCTACGATTCCTG 302  
Qy  
181 GAGACGCTGGGCAAGGCGACCTACGGGAAGGTGAAGAGGACGAGAGAGCTCGGGGCGT 240  
Db  
303 GAGACGCTGGGCAAGGCGACCTACGGGAAGGTGAAGAGGACGAGAGAGCTCGGGGCGT 362  
Qy  
241 CTGGTGCCATCAAGTCCATCAGGAAAGACAAATCAAAAGATGAGCAGGATCTGCTGCAC 300  
Db  
363 CTGGTGCCATCAAGTCCATCAGGAAAGACAAATCAAAAGATGAGCAGGATCTGCTGCAC 422  
Qy  
301 ATACGGAGGAGATGAGATCATGTCTTCACTCAACACCCCAACATCATTTGCCATCAT 360  
Db  
423 ATACGGAGGAGATGAGATCATGTCTTCACTCAACACCCCAACATCATTTGCCATCAT 482  
Qy  
361 GAAGTGTTCAGAAATAGCAGCAAGATTTGTATTTGTATGAGTATGCGAGCCGAGGGGAT 420  
Db  
483 GAAGTGTTCAGAAATAGCAGCAAGATTTGTATTTGTATGAGTATGCGAGCCGAGGGGAT 542  
Qy  
421 CTGTATGATTAATCATAGTGAAGGCGCACCGGTGAGTGAAGCGGACGCGAGCATTTCTTC 480  
Db  
543 CTGTATGATTAATCATAGTGAAGGCGCACCGGTGAGTGAAGCGGACGCGAGCATTTCTTC 602  
Qy  
481 CGACAGATCGTGTCTGCGCTGCACTACTGCAACAGCAAGCGGATCGTTCAACGAGATCTC 540  
Db  
603 CGACAGATCGTGTCTGCGCTGCACTACTGCAACAGCAAGCGGATCGTTCAACGAGATCTC 662  
Qy  
541 AAGCTGAAACATCTCTTAGATGCAATGGAACATCAAGATTTGTGATTTGGGCTC 600  
Db  
663 AAGCTGAAACATCTCTTAGATGCAATGGAACATCAAGATTTGTGATTTGGGCTC 722  
Qy  
601 TCACCTGTATACAAAGGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTCTCTACGCC 660  
Db  
723 TCACCTGTATACAAAGGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTCTCTACGCC 782  
Qy  
661 TCACCTGTATACAAAGGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTCTCTACGCC 720  
Db  
783 TCACCTGTATACAAAGGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTCTCTACGCC 842  
Qy  
721 GCGGTTCTCTGTATACCTCTGTGTCATGCGACCATGCGCTTTGACGGGCAAGGATCAATA 780  
Db  
843 GCGGTTCTCTGTATACCTCTGTGTCATGCGACCATGCGCTTTGACGGGCAAGGATCAATA 902  
Qy  
781 ACATGTTGAAGCAATCAGTAACGCGGCTTACCGTGAGCGCGCCCAAGCGCTCGATGCC 840  
Db  
903 ACATGTTGAAGCAATCAGTAACGCGGCTTACCGTGAGCGCGCCCAAGCGCTCGATGCC 962  
Qy  
841 TGTGGCTGTATCCGGTGGCTGTTAATGTTGTAACCCACCGCTCGGGCCACACTGGAGGAT 900  
Db  
963 TGTGGCTGTATCCGGTGGCTGTTAATGTTGTAACCCACCGCTCGGGCCACACTGGAGGAT 1022  
Qy  
901 GTAGCCAGTCAITGTTGGGTCAACTGGGGTTTACACACCGGAGTCGGGGCAAGGAGGCC 960  
Db  
1023 GTAGCCAGTCAITGTTGGGTCAACTGGGGTTTACACACCGGAGTCGGGGCAAGGAGGCC 1082  
Qy  
961 CTGCGTGAGGGTGGGCAACCTAGTGTGATCTTTGGCCGGGCTTCCATGGCGGATGTTA 1020  
Db  
1083 CTGCGTGAGGGTGGGCAACCTAGTGTGATCTTTGGCCGGGCTTCCATGGCGGATGTTA 1142  
Qy  
1021 CGTGGCTCTCGCGCCCTCTCTGGAGATGGAGCCAAAGTGTGACGCTTCTTCAAGCAG 1080  
Db  
1143 CGTGGCTCTCGCGCCCTCTCTGGAGATGGAGCCAAAGTGTGACGCTTCTTCAAGCAG 1202  
Qy  
1081 CAGTGCGGGAGGTGAAGCATGTATCTGGGCTGGAGCGGCAACATTTCTTTAAGAAG 1140  
Db  
1203 CAGTGCGGGAGGTGAAGCATGTATCTGGGCTGGAGCGGCAACATTTCTTTAAGAAG 1262  
Qy  
1141 TCCGGAAGGAGATGATGCTCAAAATCTGCAAGTGACCGCGGTGAGGATACCTCT 1200  
Db  
1263 TCCGGAAGGAGATGATGCTCAAAATCTGCAAGTGACCGCGGTGAGGATACCTCT 1322

1201 TCTCGCCTCGGCAAGAGCAGCCTTAAGCTTTCCGAAGGCAATTTCTCAAGAAAGTCTCTCT 1260  
Db  
1323 TCTCGCCTCGGCAAGAGCAGCCTTAAGCTTTCCGAAGGCAATTTCTCAAGAAAGTCTCTCT 1382  
Qy  
1261 ACCTCGTCAGGGAGGTACAGGAGGACCCCTCAGGAACTCAGACCGGTGCTGATCTCCA 1320  
Db  
1383 ACCTCGTCAGGGAGGTACAGGAGGACCCCTCAGGAACTCAGACCGGTGCTGATCTCCA 1442  
Qy  
1321 GGGCAGCTGCTCCCTGCTGTATCCCTGCTCCAGGAAAGGCACTCTTAAGAGTCTCGA 1380  
Db  
1443 GGGCAGCTGCTCCCTGCTGTATCCCTGCTCCAGGAAAGGCACTCTTAAGAGTCTCGA 1502  
Qy  
1381 CAGCGTAATCTGCTTACTACTCTCTCAGAGCCAGCGAGTCTGGGAACTCTTAGAC 1440  
Db  
1503 CAGCGTAATCTGCTTACTACTCTCTCAGAGCCAGCGAGTCTGGGAACTCTTAGAC 1562  
Qy  
1441 GCCAGTGATGTGTTGTGAGTGGGACCCCGTGGAGCAGAAAGTCTCCACAGGCTTCAGGG 1500  
Db  
1563 GCCAGTGATGTGTTGTGAGTGGGACCCCGTGGAGCAGAAAGTCTCCACAGGCTTCAGGG 1622  
Qy  
1501 CTCTCTCTCAACGCAAGGCAATTCATACTCAATGGCAAGTCTTCCCGCAGAGCCTTA 1560  
Db  
1623 CTCTCTCTCAACGCAAGGCAATTCATACTCAATGGCAAGTCTTCCCGCAGAGCCTTA 1682  
Qy  
1561 GAAGGCACTACCCCTAGCAGCTTTGGCTCCCTGACCACTGGCCTCTCCCATCTGCA 1620  
Db  
1683 GAAGGCACTACCCCTAGCAGCTTTGGCTCCCTGACCACTGGCCTCTCCCATCTGCA 1742  
Qy  
1621 GCCCGGCCAGCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTCTGCTCCGAGTCC 1680  
Db  
1743 GCCCGGCCAGCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTCTGCTCCGAGTCC 1802  
Qy  
1681 TTTGACCAATTTGATGCTGCTGAACTTTCCGAAACCCCACTAGGGGCTGTGTGTCT 1740  
Db  
1803 TTTGACCAATTTGATGCTGCTGAACTTTCCGAAACCCCACTAGGGGCTGTGTGTCT 1862  
Qy  
1741 GTGCAACACTGAGGGGCTTTGAGCAGCTCCCTCAGAGGCTCTGAAGCGATGTTGGCAG 1800  
Db  
1863 GTGCAACACTGAGGGGCTTTGAGCAGCTCCCTCAGAGGCTCTGAAGCGATGTTGGCAG 1922  
Qy  
1801 GAATCTCTGGGGATAGTCTCTTCTGACAGACTGCCAAGAGGTGATGTCAGGCTAC 1860  
Db  
1923 GAATCTCTGGGGATAGTCTCTTCTGACAGACTGCCAAGAGGTGATGTCAGGCTAC 1982  
Qy  
1861 AGACAGCCCTAGGAATCTGCTCAAGGCTCAGC 1893  
Db  
1983 AGACAGCCCTAGGAATCTGCTCAAGGCTCAGC 2015

## RESULT 2

US-10-322-281-22  
; Sequence 22, Application US/10322281  
; Publication No. US20040126762A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: 52945201000  
; CURRENT APPLICATION NUMBER: US/10/322,281  
; CURRENT FILING DATE: 2002-12-17  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 3073  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-322-281-22

Query Match 100.0%; Score 1893; DB 17; Length 3073;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGTTCGGTGGCTTACTTCCAGCGCCGAGCAGGCTCCCTCGGCTCCGCGCTCGCC 60



Db 92 ATGAGTCGCTGGCTTACTCCAGCGCCGAGCAGGCTCCCTCGGCTCCGCTCGGCC 151  
Qy 61 TCGGAGAGCGCCCGCTGGCGGACGGCTCATCAAGTCGCCCTAAACCTCTTGATGAAG 120  
Db 152 TCGGAGAGCGCCCGCTGGCGGACGGCTCATCAAGTCGCCCTAAACCTCTTGATGAAG 211  
Qy 121 AAGCAGCGGTGAAGCGGCAACCATCAAAACACAACTCGGCGCACCGCTACGAGTTCCTG 180  
Db 212 AAGCAGCGGTGAAGCGGCAACCATCAAAACACAACTCGGCGCACCGCTACGAGTTCCTG 271  
Qy 181 GAGACGCTGGGCAAGGCACTACGGGAAGGTGAAGAGGCAAGAGAGCTCGGGCGT 240  
Db 272 GAGACGCTGGGCAAGGCACTACGGGAAGGTGAAGAGGCAAGAGAGCTCGGGCGT 331  
Qy 241 CTGGTGGCCATCAAGTCCATCAGGAAAGACAAATCAAGATGAGCAGGATCTGCTGCAC 300  
Db 332 CTGGTGGCCATCAAGTCCATCAGGAAAGACAAATCAAGATGAGCAGGATCTGCTGCAC 391  
Qy 301 ATACGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAATCATTTGCCATCCAT 360  
Db 392 ATACGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAATCATTTGCCATCCAT 451  
Qy 361 GAAGTGTGAGATACAGCAAGATTGTGATGTCTATGAGTATGCGACGCGAGGCGAT 420  
Db 452 GAAGTGTGAGATACAGCAAGATTGTGATGTCTATGAGTATGCGACGCGAGGCGAT 511  
Qy 421 CTGTATGATTACATCAGTGAAGCGCCACGGCTGAGTGAGCGGAGCGCATTTCTTC 480  
Db 512 CTGTATGATTACATCAGTGAAGCGCCACGGCTGAGTGAGCGGAGCGCATTTCTTC 571  
Qy 481 CGACAGATCGTGTCTGCGCTGCACTACTGCGCACAGAAACGGGATCGTTTCAACGAGATCTC 540  
Db 572 CGACAGATCGTGTCTGCGCTGCACTACTGCGCACAGAAACGGGATCGTTTCAACGAGATCTC 631  
Qy 541 AAGCTGAAACATCTCTTAGATGCOATGGAACATCAAGATGTCTGATCTGCTCTG 720  
Db 632 AAGCTGAAACATCTCTTAGATGCOATGGAACATCAAGATGTCTGATCTGCTCTG 691  
Qy 601 TCCAACTGTACCAACAAAGCAAGTTCCTCCAGACGTTCTGTGGAGCCCTCTCTAGGCC 660  
Db 692 TCCAACTGTACCAACAAAGCAAGTTCCTCCAGACGTTCTGTGGAGCCCTCTCTAGGCC 751  
Qy 661 TCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCGAGAGTGAGAGTGTCTGCTG 720  
Db 752 TCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCGAGAGTGAGAGTGTCTGCTG 811  
Qy 721 GCGCTTCTCTGTACATCTGCTGATGCGCAACATGCCCCCTTTGACGGGCGAGGATCAATA 780  
Db 812 GCGCTTCTCTGTACATCTGCTGATGCGCAACATGCCCCCTTTGACGGGCGAGGATCAATA 871  
Qy 781 ACAGTGTGAAGCAAAATCAGTGAAGGCGCTTACCGTGAGCGCCCAAGCCGTCGATGCC 840  
Db 872 ACAGTGTGAAGCAAAATCAGTGAAGGCGCTTACCGTGAGCGCCCAAGCCGTCGATGCC 931  
Qy 841 TGTGGCTGTATCCGGTGGCTGTTAATGGTGAACCCCAACCCCGTCGGGCGCACACTGGAGAT 900  
Db 932 TGTGGCTGTATCCGGTGGCTGTTAATGGTGAACCCCAACCCCGTCGGGCGCACACTGGAGAT 991  
Qy 901 GTAGCCAGTCAATTGGTGGGTCAACTGGGGTTACACCAACCGAGTTCGGGGAACAGGAAGCC 960  
Db 992 GTAGCCAGTCAATTGGTGGGTCAACTGGGGTTACACCAACCGAGTTCGGGGAACAGGAAGCC 1051  
Qy 961 CTGGCTGAGGTTGGGCAACCTAGTGGTGAATTTGGCGGGCCCTCATGGCGGACTGGTTA 1020  
Db 1052 CTGGCTGAGGTTGGGCAACCTAGTGGTGAATTTGGCGGGCCCTCATGGCGGACTGGTTA 1111  
Qy 1021 CGTGGCTCTCGCGCCCGCTCTCTGGAGAAAGGAGGAGTGTGACGCTTCTTCAAGCAG 1080  
Db 1112 CGTGGCTCTCGCGCCCGCTCTCTGGAGAAAGGAGGAGTGTGACGCTTCTTCAAGCAG 1171  
Qy 1081 CAGTGCAGGAGGTGAAGCACTGTACTCTGGCTGGAGCGGCAACATCTCTTAAAGAG 1140

Db 1172 CACGTGCCGGAGGTGAAGACACTGTACTCTGGGCTGGAGCGGCAACATCTCTTAAAGAG 1231  
Qy 1141 TCCCGAAAGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCGGCTGAGGATACCTCT 1200  
Db 1232 TCCCGAAAGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCGGCTGAGGATACCTCT 1291  
Qy 1201 TCTCGGCTTGGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATCTCAAGAAAAAGTCTCT 1260  
Db 1292 TCTCGGCTTGGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATCTCAAGAAAAAGTCTCT 1351  
Qy 1261 ACCTCGCTCAGGAGGCTACAGGAGGACCTCAGGAACTCAGACCGGTCCTGATACCTCA 1320  
Db 1352 ACCTCGCTCAGGAGGCTACAGGAGGACCTCAGGAACTCAGACCGGTCCTGATACCTCA 1411  
Qy 1321 GGCAGCTGTCTCCCTGCTGTATCCCTGCTCCCAAGAAAGGCAATCTTAAAGAAAGTCTCGA 1380  
Db 1412 GGCAGCTGTCTCCCTGCTGTATCCCTGCTCCCAAGAAAGGCAATCTTAAAGAAAGTCTCGA 1471  
Qy 1381 CAGCGTGAATCTCGGTTACTACTCTCTCAGAGCCGAGAGTCTGGGGAACCTCTTAGAC 1440  
Db 1472 CAGCGTGAATCTCGGTTACTACTCTCTCAGAGCCGAGAGTCTGGGGAACCTCTTAGAC 1531  
Qy 1441 GCCAGTGTGTCTGTGAGTGGGACCCGCTGGAGCAGAGTCTCCACAGGCTTCAGGG 1500  
Db 1532 GCCAGTGTGTCTGTGAGTGGGACCCGCTGGAGCAGAGTCTCCACAGGCTTCAGGG 1591  
Qy 1501 CTCTCTCCACCGCAAGGCAATCTCAAACTCAATGGCAAGTCTCCCGCACAGCCTTA 1560  
Db 1592 CTCTCTCCACCGCAAGGCAATCTCAAACTCAATGGCAAGTCTCCCGCACAGCCTTA 1651  
Qy 1561 GAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCTCTCTCCCATCTGCA 1620  
Db 1652 GAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCTCTCTCCCATCTGCA 1711  
Qy 1621 GCCCGGCCACCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTGTCTCGAGTCC 1680  
Db 1712 GCCCGGCCACCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTGTCTCGAGTCC 1771  
Qy 1681 TTTGACCAATTTGGACTTTGGCTGAACCTTTCCCGAAACCCCACTGAGGGGCTGTGTCT 1740  
Db 1772 TTTGACCAATTTGGACTTTGGCTGAACCTTTCCCGAAACCCCACTGAGGGGCTGTGTCT 1831  
Qy 1741 GTGCAACACTGAGGGGCTTGAAGCAGCTCTCTCAGAAAGTCTGAAGCGATGGTGGCAG 1800  
Db 1832 GTGCAACACTGAGGGGCTTGAAGCAGCTCTCTCAGAAAGTCTGAAGCGATGGTGGCAG 1891  
Qy 1801 GAATCTTGGGATAGTGTCTTTCTGACAGACTGCCAAGAGTGTGACGCTAC 1860  
Db 1892 GAATCTTGGGATAGTGTCTTTCTGACAGACTGCCAAGAGTGTGACGCTAC 1951  
Qy 1861 AGACAGCCCTAGGAATCTGCTCAAAAGCTCAGC 1893  
Db 1952 AGACAGCCCTAGGAATCTGCTCAAAAGCTCAGC 1984

## RESULT 3

US-10-343-514-87

; Sequence 87, Application US/10343514

; Publication No. US20040132025A1

; GENERAL INFORMATION:

; APPLICANT: DRUCKER, Daniel J.

; APPLICANT: ROSEN, Cheryl F.

; APPLICANT: LEFEBVRE, Diana L.

; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK

; FILE REFERENCE: DPA-DRUC2/PCT

; CURRENT APPLICATION NUMBER: US/10/343,514

; CURRENT FILING DATE: 2003-01-31

; PRIOR APPLICATION NUMBER: PCT/CA01/01109

; PRIOR FILING DATE: 2001-08-02

; PRIOR APPLICATION NUMBER: US 60/222,650

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: US 60/274,613

; PRIOR FILING DATE: 2001-03-12

;; PRIOR APPLICATION NUMBER: CA 2,340,780  
;; PRIOR FILING DATE: 2001-03-28  
;; NUMBER OF SEQ ID NOS: 109  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 87  
;; LENGTH: 2027  
;; TYPE: DNA  
;; ORGANISM: MOUSE  
US-10-343-514-87

Query Match 99.7%; Score 1888.2; DB 17; Length 2027;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1890; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	1	ATGAGTCGCTGGGCTTACTCCAGCGCCGAGCGAGCTCCCTCGGCTCGGCTCGCC	60
Db	75	ATGAGTCGCTGGGCTTACTCCAGCGCCGAGCGAGCTCCCTCGGCTCGGCTCGCC	134
Qy	61	TCGAGAGCGCCCGCGCTGGCGGAGCGGCTCATCAAGTCGCCTAAACCTCTGATGAAG	120
Db	135	TCGAGAGCGCCCGCGCTGGCGGAGCGGCTCATCAAGTCGCCTAAACCTCTGATGAAG	194
Qy	121	AAGCAGCGGTGAAGCGGACCATCAAAACAACTCTGGGCAACCGGTACGATTCCTG	180
Db	195	AAGCAGCGGTGAAGCGGACCATCAAAACAACTCTGGGCAACCGGTACGATTCCTG	254
Qy	181	GAGACGCTGGGCAAGGCGCCTACGGAAGGTGAAGAGCGACGAGAGCTCGGGCGT	240
Db	255	GAGACGCTGGGCAAGGCGCCTACGGAAGGTGAAGAGCGACGAGAGCTCGGGCGT	314
Qy	241	CTGCTGCCCATCAAGTCCATCAGAAAGACAAATCAAGATGAGAGGATCTGCTGCAC	300
Db	315	CTGCTGCCCATCAAGTCCATCAGAAAGACAAATCAAGATGAGAGGATCTGCTGCAC	374
Qy	301	ATACGAGGAGATTGATCATGTCTTCACTCAACCAACCCCAATCATTCGATCCAT	360
Db	375	ATAAGGAGGAGATTCGAGATCATGTCTTCACTCAACCAACCCCAATCATTCGATCCAT	434
Qy	361	GAAGTGTTCGAGATACGACAGATTGTGATGTCATGAGTATGCGACCGAGCGGAT	420
Db	435	GAAGTGTTCGAGATACGACAGATTGTGATGTCATGAGTATGCGACCGAGCGGAT	494
Qy	421	CTGTATGATTACATCAGTGAAGCGCCACGGCTGAGTGAGGCGGACGCGAGGCAATTCCTC	480
Db	495	CTGTATGATTACATCAGTGAAGCGCCACGGCTGAGTGAGGCGGACGCGAGGCAATTCCTC	554
Qy	481	CGACAGATCGTGTCTGCCCTGCACTACTGCCACAGAAACGGGATCGTTTCAACGAGATCTC	540
Db	555	CGACAGATCGTGTCTGCCCTGCACTACTGCCACAGAAACGGGATCGTTTCAACGAGATCTC	614
Qy	541	AAGCTGAAACATCTCTTAGATGCCAATGGAACATCAAGATTGCTGACTTTGGGCTC	600
Db	615	AAGCTGAAACATCTCTTAGATGCCAATGGAACATCAAGATTGCTGACTTTGGGCTC	674
Qy	601	TCCAACCTGTACCAAAAGGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTCTCTAGGCC	660
Db	675	TCCAACCTGTACCAAAAGGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTCTCTAGGCC	734
Qy	661	TCGCTGAGATAGTCAACGGGAGCCCTATGTGGGCCAGAGGTGGAACAGTGTCTCTG	720
Db	735	TCGCTGAGATAGTCAACGGGAGCCCTATGTGGGCCAGAGGTGGAACAGTGTCTCTG	794
Qy	721	GGCGTCTCTGTATACCTCTGTCATGGCACCATGCCCTTTGACGGGAGGATCATAAA	780
Db	795	GGCGTCTCTGTATACCTCTGTCATGGCACCATGCCCTTTGACGGGAGGATCATAAA	854
Qy	781	ACACTGTGTAAGCAAAATCAGTAAACGGGCTTACCGTGAGCGGCCCAAGCCGTCGATGCC	840
Db	855	ACACTGTGTAAGCAAAATCAGTAAACGGGCTTACCGTGAGCGGCCCAAGCCGTCGATGCC	914
Qy	841	TGTGGCCCTGATCCGGTGGCTGTTAATGGTGAACCCCAACCGGTGGGGCCACACTGGAGGAT	900
Db	915	TGTGGCCCTGATCCGGTGGCTGTTAATGGTGAACCCCAACCGGTGGGGCCACACTGGAGGAT	974

RESULT 4

US-10-343-514-27

; Sequence 27, Application US/10343514

Publication No. US20040132025A1  
GENERAL INFORMATION:  
APPLICANT: DRUCKER, Daniel J.  
APPLICANT: ROSEN, Cheryl F.  
APPLICANT: LEBEVRE, Diana L.  
TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK  
FILE REFERENCE: DPA-DRUG2/PCT  
CURRENT APPLICATION NUMBER: US/10/343,514  
CURRENT FILING DATE: 2003-01-31  
PRIORITY APPLICATION NUMBER: PCT/CA01/01109  
PRIORITY FILING DATE: 2001-08-02  
PRIORITY FILING DATE: 2000-08-03  
PRIORITY FILING DATE: 2000-08-03  
PRIORITY FILING DATE: 2001-03-12  
PRIORITY FILING DATE: 2001-03-12  
PRIORITY FILING DATE: 2001-03-28  
NUMBER OF SEQ ID NOS: 109  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 27  
LENGTH: 2026  
TYPE: DNA  
ORGANISM: RAT  
US-10-343-514-27

Query Match 88.9%; Score 1683.4; DB 17; Length 2026;  
Best Local Similarity 93.4%; Pred. No. 0;  
Matches 1769; Conservative 0; Mismatches 121; Indels 3; Gaps 1;  
QY 1 ATGAGTCGCTGGCTTACTCCAGCGCCGAGCCGAGCGCTCCCTCGCGCTCCGCGCTGGCC 60  
DB 75 ATGAGTCGCTGGCTTACTACCGCGCGGGAACCTGGCTCCCTCGCGCTCCGCGCTGGCC 134  
QY 61 TCGAGAGCGCCGCGCTGGCGGCGGCTCATCAAGTCGCTAAACCTCTGATGAAG 120  
DB 135 ACGAGAGCGCCGCGCTGGCGGCGGCTCATCAAGTCGCGGCAACCTCTGATGAAG 194  
QY 121 AAGCAGCGCTGAAGCGGCACCATCAAAACAAACCTCGCGGCAACCGCTACGAGTTCCTG 180  
DB 195 AAGCAGCGCTGAAGCGGCACCATCAAAACAAACCTCGAGGCAACCGCTACGAGTTCCTG 254  
QY 181 GAGACGCTGGGCAAGGCACTACGGAAGGTGAAGAGGCAAGAGAGCTCGGGCGGT 240  
DB 255 GAGACGCTGGGCAAGGCACTACGGAAGGTGAAGAGGCAAGAGAGCTCGGGACGC 314  
QY 241 CTGCTGCCATCAAGTCCATCAGAAAGACAAATCAAGATGAGCAGGATCTGCTGCAC 300  
DB 315 CTGCTGCCATCAAGTCTATCAGAAAGCAAAATCAAGATGAGCAGGATCTGTTGCAC 374  
QY 301 ATACGAGGAGGATTCAGATCATGCTTCTCACTCAACCCACCCACATCATTTGCCATCCAT 360  
DB 375 ATAAGAGGAGGATTCAGATCATGCTTCTCACTCAACCCACCCACATCATTTGCCATCCAT 434  
QY 361 GAAGTGTTCAGAAATAGCAGCAAGATTCGATTCATGATGATGATGATGATGATGATGATGAT 420  
DB 435 GAAGTGTTCAGAAATAGCAGCAAGATTCGATTCATGATGATGATGATGATGATGATGATGAT 494  
QY 421 CTGATGATTCATCAGTACGCGGCAAGCTGAGTACGCGGCAAGCTGAGTACGCGGCAAGCTGAG 480  
DB 495 CTGATGATTCATCAGTACGCGGCAAGCTGAGTACGCGGCAAGCTGAGTACGCGGCAAGCTGAG 554  
QY 481 CGACAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
DB 555 CGACAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 614  
QY 541 AAGCTGGAACAAATCTCTTCAGATGCAATGGAACAAATGGAACAAATGGAACAAATGGAACAAAT 600  
DB 615 AAGCTGGAACAAATCTCTTCAGATGCAATGGAACAAATGGAACAAATGGAACAAATGGAACAAAT 674  
QY 601 TCCAACTGTACCAAGGCAAGTTCCTCCAGACGTTCTGTTGGGAGCCCTCTCTAGGCC 660  
DB 675 TCCAACTGTATCACAAAGGCAAGTTCCTCCAGACGTTCTGTTGGGAGCCCTCTCTATGCC 734

QY 661 TCGCTCAGATAGTCAACGGGAAGCCCTATGTGGGCCAGAGGTGGACAGCTGGTCTCTG 720  
DB 735 TCACTCAGATCGTCAACGGGAAGCCCTATGTGGGCCAGAGGTGGACAGCTGGTCTCTG 794  
QY 721 GGGTTCCTCCTGATACATCCTGGTGCATGGCAACATGCTTTGACGGGCAAGATCATATAA 780  
DB 795 GGGTTCCTCCTGATACATCCTGGTGCATGGCAACATGCTTTGACGGGCAAGATCATATAA 854  
QY 781 ACATCGTGAAGCAAAATCAGTAACGGGCTTACCGTACGCGCCCAAGCCGTCGATGCC 840  
DB 855 ACCCTGTGAACAAATCAGTAGGGGCTTACCGAGCGCGTGAACACCGCTGATGCC 914  
QY 841 TGTGGCCTGATCCGGTGGCTGTTAATGGTGAACCCCAACCGTGGGGCACACATGGAGAT 900  
DB 915 TGTGGCCTGATCCGGTGGCTGTTAATGGTGAACCCCAACCGTGGGGCACATCTGGAGAT 974  
QY 901 GTAGCCAGTCAATTTGGTGGTCAACTGGGGTTACACACCGGAGTTCGGGGAACAGGAAGCC 960  
DB 975 GTAGCCAGTCAATTTGGTGGTCAACTGGGGTTACAGCACCCGCAATTTGGGGAACAGGAAGCT 1034  
QY 961 CTGCTGAGGCTGGGCAACCTAGTGGTGAATTTGGCGGGCCCTCCATGGCGGACTGGTTA 1020  
DB 1035 CTGCGAGAGGCTGGGCAACCTAGTGGTGAATTTGGCGGGCCCTCTATGGCGGACTGGTTA 1094  
QY 1021 CGTCTGCTCTCGCGCCCTCTCTCGAGAATGGAGCCAAAGGTGTCAGCTTCTTCAAGCAG 1080  
DB 1095 CGTCTGCTCTCGCGCCCTCTCTCGAGAATGGAGCCAAAGGTGTCAGCTTCTTCAAGCAG 1154  
QY 1081 CAGTTCGGGAGGTGAAGCACTGTACTGGGTGAGCGGCAACATTTCTTTAAGAAG 1140  
DB 1155 CATGTGCGGAGGTGAAGCACTGTACTGGGTGAGCGGCAACATTTCTTTAAGAAG 1214  
QY 1141 TCCGGAAGGAGAAATGACATGGCTCAAAATCTCAAGGTGACCGGCTGAGGATACCTCT 1200  
DB 1215 TCCGGAAGGAGAAATGACATGGCTCAAGTCTCAGACTCTCAGAAATGACCGAGTTGAAGTACTTCC 1274  
QY 1201 TCTCGCCCTGGCAAGAGCAGCCCTTAAGCTTTCGAAAGGCAATTTCTCAAGAAAGGCTCTCT 1260  
DB 1275 TCTCGCCCTGGCAAGAGCAGCCCTCAAGCTTTCGAAAGGCAATTTCTCAAGAAAGGCTCTCT 1334  
QY 1261 ACCTGTCAGGAGGTGACAGAGGACCTCTCAGAACTCAGACCGGTGCTGATCTCCA 1320  
DB 1335 CCCTCATCGGAGGAGGTACAGAGGCGCCTCAGAACTCAGACGAGTGTCCAAATACCCCA 1394  
QY 1321 GGGCAGCTGCTCCCTGCTGATCCCTGCTCCCAAGGAAGGCAATCTTTAAGAGTCTCGA 1380  
DB 1395 GGGCAGCTGCTCCCTGCTGATCCCTGCTCCCAAGGAAGGCAATCTTTAAGAGTCTCGG 1454  
QY 1381 CAGCGTGAATCTGGTTACTACTCTCTCCAGAGCCCGAGTCTGGGGAACTCTTAGAC 1440  
DB 1455 CAGCGTGAATCTGGTTACTACTCTCTCCAGAGCCCGAGTCTGGGAACTCTTAGAC 1514  
QY 1441 GCAGTGTATGTTTGTGAGTGGGACCCCGTGGAGAGAAAGTCTCCACAGGCTTCAGGG 1500  
DB 1515 GCAGTGTATGTTTGTGAGTGGGACCCCGTGGAGAGAAAGTCTCCACAGGCTTCAGGG 1574  
QY 1501 CTCTCTCCACCGCAAGGCAATTTCTCAAACTCAATGGCAAGTTCCTCCCGCAGAGCCTTA 1560  
DB 1575 ---CGCTCTCATCGAAGGCAATTTCTCAAACTCAATGGCAAGTTCCTCCCGCAGAGCCTTA 1631  
QY 1561 GAAGGCACTACCCCTAGCAGCTTTGGCTCCCTGGACCAACTGGGCTCTCTCCCATCTTGCA 1620  
DB 1632 GAAGGCACTGCCCCCTAGCAGCTTTGGCTCCCTAGACCAACTGGGCTCTCTCATCTACA 1691  
QY 1621 GCCCGCCAGCCGCCCTCAGGGGCTGTGAGTGAAGAGCAGCATCTCTGCTCCGAGTCC 1680  
DB 1692 GCCCGCCAGCCGCCCTCAGGGAGCTGTGAGTGAAGAGCAGCATCTCTGCTCCGAGTCC 1751  
QY 1681 TTTGACCAATTTGACCTGCTGAACTCTTCCCGAAGCCCACTGAGGGGCTGTGTGCT 1740  
DB 1752 TTTGACCAATTTGACCTGCTGAACTCTTCCCGAGCGGCTTCCCGAAGCCCACTGAGGAGCTGTGTGCT 1811  
QY 1741 GTGGCAACCTGAGGGGGCTTTGAGCAGGCTCTCCCTCAGAAAGGTCTGAAGCGATGTTGGCAG 1800



Db 1583 ---GCGCTCCANTCGAAGGGGATCCTCAAACTCAATGGCAAGTTTCCCCACAGCCCTTA 1639  
Qy 1561 GAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGGCTCTCTCCATCCTGCA 1620  
Db 1640 GAAGGCACTGCCCTTAGCACCTTTGGCTCCCTAGACCAACTGGGCTCCCTCATCTTACA 1699  
Qy 1621 GCGCGGCCAGCCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTCTCTCCGAGTCC 1680  
Db 1700 GCGCGGCCAGCCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTCTCTCCGAGTCC 1759  
Qy 1681 TTGTACCAATTTGACTTGGCTTGAACGCTTTCCGCAAAACCCCACTGAGGGCTGTGTCT 1740  
Db 1760 TTGTACCAATTTGACTTGGCTTGAACGCTTTCCGCAAAACCCCACTGAGGAGCTGTGTCT 1819  
Qy 1741 GTGACAACTTGAGGGGCTTTGACAGCTCCCTCAGAAGTCTGAAGCGATGTGGCAG 1800  
Db 1820 GTGACAACTTGAGGAGCTTTGACAGCTCCCTCAGAGGCTTAAACGATGTGGCAG 1879  
Qy 1801 GAATCCTTTGGGGATAGTCTGTTTTCTTGACAGACTGCCAAGAGGTGACTGACGCTTAC 1860  
Db 1880 GAATCCTTTGGGGATAGTCTGTTTTCTTGACAGACTGCCAAGAGGTGACTGACGAGCTTAC 1939  
Qy 1861 AGCAAGCCCTAGGAATCTGCTCAAGCTCAGC 1893  
Db 1940 AGCAAGCCCTAGGAATCTGCTGAAGCTCAGC 1972

RESULT 6  
US-09-963-159-3  
; Sequence 3, Application US/09963159  
; Patent No. US2002007312A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Galvin, Katherine M.  
; TITLE OF INVENTION: 3700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR  
; FILE REFERENCE: 10147-5001  
; CURRENT APPLICATION NUMBER: US/09/963,159  
; CURRENT FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 60/234,922  
; PRIOR FILING DATE: 2000-09-25  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 1884  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-963-159-3

Query Match  
Best Local Similarity 84.2%; Pred. No. 0;  
Matches 1565; Conservative 0; Mismatches 273; Indels 21; Gaps 2;  
Qy 47 CTTCCGCCCTGGCTCGGAGCGCCCGCGCTGGCGGAGGGCTCATCAAGTCGCTTA 106  
Db 35 CCATCCCTCGCGCGAGACTAGCCCGCGCTGGCGGAGGGCTGATCAAGTCGCGCA 94  
Qy 107 AACCTCTGATGAAGACAGCGGTGAAGCGGACCATCAAAACACAACTCGCGCAC 166  
Db 95 AGCCCTTAATGAAGACAGCGGTGAAGCGGACCATCAAAACACAACTCGCGCAC 154  
Qy 167 GCTACGAGTTCTGGAGACGCTGGGCAAGGCGACCTACGGGAAGGTGAAGAGGACGAG 226  
Db 155 GCTACGAGTTCTGGAGACGCTGGGCAAGGCGACCTACGGGAAGGTGAAGAGGCGCGG 214  
Qy 227 AGAGCTCGGGGCTGTGGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGC 286  
Db 215 AGAGCTCGGGGCGCTGTGGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGC 274  
Qy 287 AGATCTGTGCATACAGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCA 346  
Db 275 AGATCTGTGCATACAGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCA 334  
Qy 347 TCATTGCCATCCATGAAGTGTGTTGAGATAGACGAAGATTGTGATGTCAGGATG 406

Db 335 TCATTGCCATCCATGAAGTGTGTTGAGAAAGCAGCAAGATCGTGATCGTCATGAGTATG 394  
Qy 407 CGAGCGAGGGATCTGTATGATTAATCATCATGAGCGGCGCACGGCTGAGTGAGCGGAGC 466  
Db 395 CGAGCGGGGCGACCTTTATGACTACATCAGCAGCGGCGGACAGCTCAGTGAGCGGAG 454  
Qy 467 CCAGGCAATTTCTCCGACAGATCGTGTCTGCCCTGCACTACTGCCACAGAACCGGATCG 526  
Db 455 CTAGGCATTTCTCCGCGAGATCGTCTCTGCCGTGCACTATTGCCATCAGAACAGATTG 514  
Qy 527 TTCAACGAGATCTCAAGCTCGAAACATCTCTTCTAGATGCCAATGGAACATCAAGATTG 586  
Db 515 TCACACGAGATCTCAAGCTGGAGAACATCTCTTTGGATGCCAATGGAACATCAAGATTG 574  
Qy 587 CTGACTTTGGGCTCTCCAGCTCTACCAACAGCAAGTTCTTCCACAGCTTCTGTGGGA 646  
Db 575 CTGACTTTGGGCTCTCCAGCTCTTACCATCAAGCAAGTTCTTCCACAGCATTTCTGTGGGA 634  
Qy 647 GCCCTCTACGCGCTCGCCTGAGATAGTCAACGGAAGCCCTATGTGGGCCAGAGTGG 706  
Db 635 GCCCTCTATGCTCGCCAGAGATTGTCAATGGGAAGCCCTACACAGGCCAGAGTGG 694  
Qy 707 ACAGCTGGTCTCTGGGCGTTCTCTCTGATCTCTGTGTCATGTCATGTCACCATGTCGCG 766  
Db 695 ACAGCTGGTCTCTGGGCGTTCTCTCTGATCTCTGTCATGTCGTCATGTCGTCATGTCG 754  
Qy 767 GCGAGGATCAATAAANACTGGTGAAGCAATCAGTAACGGGGCTTACCCTGAGCGGCCCA 826  
Db 755 GCGATGACCATTAAGATCTTAGTGAACACAGATCAGCAACGCGGGCTTACCGGAGCACCTA 814  
Qy 827 AGCGCTCCGATGCTGTGGGCTCATCGGTGGCTGTTAATGGTGAACCCACCCGTCGGG 886  
Db 815 AACCTCTGATGCTGTGGCTCATCGGTGGCTGTTGATGGTGAACCCACCCGTCGGG 874  
Qy 887 CCACTCTGAGGATGTAGCCAGTCACTTGGTGGGTCAACTGGGGTTACACCAACCGAGTGG 946  
Db 875 CCACCTCGAGGATGTGGCCAGTCACTGGTGGGTCAACTGGGGTTACACCAACCGAGTGG 934  
Qy 947 GGGAAAGGNAAGCCCTCGCTGAGGGTGGGCAACCTAGTGGTGAATTTGGCGGCGCTCCA 1006  
Db 935 GAGAGCAGGAGGCTCCGCAATGAGGGTGGGCAACCTGSCAGTGACTCTGCGCGCGCTCCA 994  
Qy 1007 TGGCGGACTGTTTACGTGCGCTCTCGCGCCCTCTCGGAGAAATGAGACCAAGGTGTGCA 1066  
Db 995 TGGCTGACTGGCTCCGGCGTTCTCTCGCGCCCTCTCTGAGAAATGAGGCGCCAGGTGTGCA 1054  
Qy 1067 GCTTCTTCAAGCAGCAGTGCCTCGGAGGTGGAAGCACTGTACTCTGGGCTGGAGCGCAAC 1126  
Db 1055 GCTTCTTCAAGCAGCAGTGCCTCGTGGGGGAAGCACCCCTTGGCTTGGAGCGCCAGC 1114  
Qy 1127 ATTCTCTTAAGAAAGTCCGNAAGGAGATGACATGGCTCAAAATCTGCAAGGTGACCCGG 1186  
Db 1115 ATTCTCTTAAGAAAGTCCGNAAGGAGATGACATGGCTCAAAATCTGCAAGGTGACCCGG 1174  
Qy 1187 CTGAGGATACCTCTCTCGGCTTGGCAAGCAGCCTTAAGCTTCCGAAAGGCAATCTCA 1246  
Db 1175 CTGATGACACTGCGCATCGGCTTGGCAAGGAACTCAAGCTGCGCAAGGCAATCTCA 1234  
Qy 1247 AGAAAGTCTCTTACCTCTCGTCAAGGAGGTACAGGAGGACCTCAGGAACTCAGACCGG 1306  
Db 1235 AGAAAGGTTGTGAGCTCTGCAAGAGGGGTACAGGAGGACCTCAGGAGCTCAGGCCAA 1294  
Qy 1307 TGCTGATCTCCAGGCGACCTGTCCCTGCTGTATCTCTGCTCCAGGAAAGGCAATCC 1366  
Db 1295 TCCCTCGAGCCCGAGGCGAGGCTGCCCC-----CCCTGCTCCCCAAGAGGGCAATC 1345  
Qy 1367 TTAAGAAAGTCTCGACAGCGTGAATCTGTTTACTACTCTCTCCAGAGCCCGAGGAGTCTG 1426  
Db 1346 TCAAGAAAGCCCGACAGCGGAGTCTGCTACTACTCTCTCCGAGCCCGAGTGAATCTG 1405  
Qy 1427 GGAACCTCTTAGACCGCAGTGATGTGTTGTGAGTGGGGACCCCGCTGGAGCAGAGTCTC 1486





```
Db 875 CCACCCCTGGAGATGTGGCCAGTCACTGGTGGGTCAACTGGGGCTAGCCACCCGAGTGG 934
Qy 947 GGAACAAGGAGCCCTCGGTGAGGGTGGGACCCCTAGTGGTGAATTTGGCGGCGCTTCCA 1006
Db 935 GAGAGCAGGAGGCTCCGATGAGGGTGGGACCCCTGCGAGTACTCTGCCGCGCCCTCCA 994
Qy 1007 TGGCGGACTGTTAGTGTCTCTCGGCGCCCTCTCGGAGAAATGGAGCCCAAGGTGTGCA 1066
Db 995 TGGCTGACTGGCTCCGGGCTTCTCTCCGCGCCCTCTGGAGAAATGGGCGCAAGGTGTGCA 1054
Qy 1067 GCTCTTTCAAGCAGCAGTGGCGGAGGTGGAAGCACTGTACCTGGGCTGGAGCGGCAAC 1126
Db 1055 GCTTCTTTCAAGCAGCAGTGGCGGAGGTGGAAGCACTGTACCTGGGCTGGAGCGGCAAC 1114
Qy 1127 ATTCTCTTTAAGAGTCCCGAAAGAGAAATGACATGGCTCAAAATCTCAAGAGTCAACCCG 1186
Db 1115 ATTCTCTTTAAGAGTCCCGAAAGAGAAATGACATGGCTCAAAATCTCAAGAGTCAACCCG 1174
Qy 1187 CTGAGGATACTCTTTCTCGCCCTGGCAAGAGCAGCTTAAAGCTTTCCGAAAGGCAATTTCA 1246
Db 1175 CTGATGACACTGGCCATCGCCCTGGCAAGAGCAGCTTAAAGCTTTCCGAAAGGCAATTTCA 1234
Qy 1247 AGAAAGAGTCTTACCTCTGTCAGGAGGAGGTACAGGAGGACCTCTAGGAACTCAGACCGG 1306
Db 1235 AGAAAGAGTCTTACCTCTGTCAGGAGGAGGTACAGGAGGACCTCTAGGAACTCAGACCGG 1294
Qy 1307 TGCTGTATCTTCCAGGAGCAGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1366
Db 1295 TCCCTGCGAGCCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1345
Qy 1367 TTAAGAGTCTCGACAGCAGTGAATCTGGTTACTACTCTCTCTCAGAGCCAGCAGGCTG 1426
Db 1346 TCAAGAGCCCGGACAGCGGAGTCTGGCTACTACTCTCTCTCAGAGCCAGCAGGCTG 1405
Qy 1427 GGGAACTTTAGAGCCAGTGAATGTGTTGAGTGGGACCCCGTGGAGCAGAAAGTCTC 1486
Db 1406 GGGAGCTTTGGAGCGCAGGCGAGCTGTTGAGTGGGAGTCCCAAGAGGAGCAGAAAGCTC 1465
Qy 1487 CACAGGCTTCAGGCTCTCTCTCCAGCGAGGCAATTTCTCAAACTCAATGGCAAGTTCT 1546
Db 1466 CGCAAGCTTCAGGCTCTCTCTCTCCAGCGAGGCAATTTCTCAAACTCAATGGCAAGTTCT 1525
Qy 1547 CCGCAGCAGCTTTAGAGGCACTACCCCTAGCAGCTTTGGCTCTCTCGAGCAACTGGGCT 1606
Db 1526 CCGCAGCAGCTTTAGAGCTCGCGGCCCCCAGCAGCTTTGGCTCTCTCGAGCAACTGGGCT 1585
Qy 1607 CTTCCCATCTGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1666
Db 1586 CACCTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1645
Qy 1667 TGTCTCCGAGTCTTTGACCAATTTGACCTTGGCTGAGCAGCTCTTCCGAAACCCCACTGA 1726
Db 1646 TGTCTCTGAGTCTTTGACCAATTTGACCTTGGCTGAGCAGCTCTTCCGAAACCCCACTGC 1705
Qy 1727 GGGGCTGTGTGTCTGTGGCAAACTCGAGGGGCTTGGAGGCGCTCTTCCGAAACCCCACTGA 1780
Db 1706 GGGGCTGTGTGTCTGTGGCAAACTCGAGGGGCTTGGAGGCGCTCTTCCGAAACCCCACTGC 1765
Qy 1781 -----GTTCTGAGCGATGTGGCAGGAATCTTTGGGGGATAGTGTCTTTCTGTGACAG 1834
Db 1766 GAAAGCTCCTGAGGCGCTGGCGGAGGATCTTTGGGGGAGAGTGTCTTTCTGACAG 1825
Qy 1835 ACTGCCAAGAGGTGACTGACGCTTACAGAGGAGGCTTAGGAATCTGCTCAAGCTCAGC 1893
Db 1826 ACTGCCAAGAGGTGACTGACGCTTACAGAGGAGGCTTAGGAATCTGCTCAAGCTCAGC 1884
```

## RESULT 8

US-09-963-159-1

; Sequence 1, Application US/09963159

; Patent No. US2002007312A1

; GENERAL INFORMATION:

```
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Galvin, Katherine M.
; TITLE OF INVENTION: 3700. A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR
; FILE REFERENCE: 10147-5001
; CURRENT APPLICATION NUMBER: US/09/963,159
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/234,922
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3353
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3268)
; NAME/KEY: unsure
; LOCATION: (3270)
; NAME/KEY: unsure
; LOCATION: (3272)
US-09-963-159-1
```

```
Query Match 71.9%; Score 1360.2; DB 9; Length 3353;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 273; Indels 21; Gaps 2;

Qy 47 CTTCCGCGCTGGCTCGGAGCGCGCGCTGGCGGAGCGGCTCATCAAGTCGCGCTA 106
Db 191 CCACTCCCTCGGCGGAGAGCTAGCCGCGCGCTGGCGGAGGCTCATCAAGTCGCGCTA 250
Qy 107 AACCTCTGTAGTGAAGAGCAGCGGCTGAAGCGGACCATCAACAACACCACTCGCGGAC 166
Db 251 AGCCCTTAATGAAGAGCAGCGGCTGAAGCGGACCATCAACAACACCACTCGCGGAC 310
Qy 167 GTACAGAGTTCTTGAGAGCCTGGGCAAGGCGCCTACGCGGAAGGTGAAGAGCGACG 226
Db 311 GCTACAGAGTTCTTGAGAGCCTGGGCAAGGCGCCTACGCGGAAGGTGAAGAGCGCGG 370
Qy 227 AGAGCTCGGGGCTCTGGTGGCCATCAAGTCCATCAAGGAAGCAAAATCAAGATGAGC 286
Db 371 AGAGCTCGGGGCTCTGGTGGCCATCAAGTCCATCAAGGAAGCAAAATCAAGATGAGC 430
Qy 287 AGGATCTGTCACATACGAGGAGGATTCAGATCATGTCTTCACTCAACACCCACACA 346
Db 431 AGATCTGATGCATACGAGGAGGATTCAGATCATGTCTTCACTCAACACCCACACA 490
Qy 347 TCATTGCGATCCATGAAGTGTGAGAAATAGCAGCAAGATTGTGATGTCTATGAGTATG 406
Db 491 TCATTGCGATCCATGAAGTGTGAGAAATAGCAGCAAGATTGTGATGTCTATGAGTATG 550
Qy 407 CCAGCGAGGCGATCTGTATGATTACATCAGTGGCGGCGGCGGCGGCGGCGGCGGCGG 466
Db 551 CCAGCGGCGGCGATCTTATGACTATACATCAGCGGCGGCGGCGGCGGCGGCGGCGG 610
Qy 467 CCAGGCAATTTCTTCGACAGATCGTGTCTGCGCTGCACTACTGCGCAGCAAGCGGATCG 526
Db 611 CTAGGCAATTTCTTCGCGAGATCGTCTCTGCGCTGCACTACTGCGCAGCAAGCGGATG 670
Qy 527 TTCACCGAGATCTCAAGCTGGAAAAATCTTCTAGATGCCAATGGAAACATCAAGATG 586
Db 671 TCACCGAGATCTCAAGCTGGAAAAATCTTCTTGGATGCCAATGGAAATATCAAGATG 730
Qy 587 CTGACTTGGCGCTCTCAACCTGTACCAAGAGGAGGCTTCTTCCAGAGCTTCTGTGGGA 646
Db 731 CTGACTTGGCGCTCTCAACCTGTACCAAGAGGAGGCTTCTTCCAGAGCTTCTGTGGGA 790
Qy 647 GCCCTCTACGCTCGCGCTGAGATAGTCAACGGGAAGCGCTTATGTGGGCGGAGGCTGG 706
Db 791 GCCCTCTTATGCTTGCCAGAGATTTGTCAATGGGAAGCGCTTACACAGGCGGAGGCTGG 850
Qy 707 ACAGCTGGTCTCTGGGCGTTCTCTGTATCTCTGTGATGCGCATGCCCATGCCCTTTGACG 766
```





[illegible]

1331	CTGATGACTCGCCATCGCCCTGGCAGAGCAACCTCAAGCTGCCAAAGGGCAATTCTCA	1399
1247	AGAAAAGTCTCTACTCGTCAGGGAGAGTACAGAGAGACCCCTCAGGAACTCAGACCGG	1306
1391	AGAAGAAGGTGTTCAGCCTCTGCAGAAAGGGGTACAGGAGGACCCCTCCGGAGCTCAGCCCAA	1450
1307	TGCCTGATACTCCAGGGCAGCTGTCCCTGCTGTATCCCTGCTCCCAAGGAAAGGCAATCC	1366
1451	TCCCTCGAGCCAGGGCAGGCTGCC-----CCCTGCTCCCAAGAGGGCAATC	1501
1367	TTAAGAAGTCTGCACAGCGTGAATCTGGTTACTCTCTCCAGAGCCCAAGGAGTCTG	1426
1502	TCAAGAAGCCCGACAGCGAGTCTGGCTACTTCTCTCCCGAGCCAGTGAATCTG	1561
1427	GGGAACTCTTAGAGCCGAGTGATGTGTTGTAGTGGGAGCCCGTGGAGCAGAAGTCTC	1486
1562	GGGAGCTCTTGGACGCGAGCGACGTGTTGTAGTGGGATCCCAAGGAGCAGAAAGCCTC	1621
1487	CACAGGCTTCAGGGCTCCTCTCCACGCAAGGGCAATCTCAAACTCAATGGCAAGTTCT	1546
1622	CGCAAGCTTCAGGGCTCCTCTCTCATCGCAAGGCATCTCAAACTCAATGGCAAGTTCT	1681
1547	CCGCGACAGCTTTAGAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCCT	1606
1682	CCCAGACAGCTTTAGAGCTCGCGGCCCAACCACTTCGGCTCCCTGGATGAACCTGCCCC	1741
1607	CCTCCCATCTGCAGCCCGCCAGCGCCCTCAGGGCTGTGAGTGAGGACAGCATCC	1666
1742	CACCTCGCCCTTCGGCCCGGCGCAGCGACCCCTCAGGGGTGTGAGCGAGGACAGCATCC	1801
1667	TGTCCTCCGAGTCTTTGACCAATTGGACTTGCCCTGAAACGTCTTCCCGAAACCCCACTGA	1726
1802	TGTCCTCTGAGTCTTTGACCACTGGACTTGCCCTGAAACGGCTCCCAAGACGCCCACTGC	1861
1727	GGGGCTGTGTCTGTGTGCAAACTGAGGGGGCTTGAGCAGCCTCCCTCAGAAG-----	1780
1862	GGGGCTGTGTGTCTGTGTGCAAACTCAGGGGGCTTGAGGAGCGCCCTCAGAGGGGCCCTG	1921
1781	-----GTCTGAAGCGATGGTGGCAGGAATCCCTTGGGGGATAGCTGCTTTCTCTGACAG	1834
1922	GAAGCTGCCTGAGGCGCTGGCGCAGGATCCTTTGGGGNACAGTCTTTTCCCTGCAG	1981
1835	ACTGCCAAGAGGTGACTGCAGCCTACAGACAAGCCCTAGGAATCTGTCTAAAGCTCAGC	1893
1982	ACTGCCAAGAGGTGACAGCGACCTACCGACAGCACTGAGGGTCTGTCTCAAGCTCACC	2040

```

RESULT 10
US-10-311-034-38
; Sequence 38, Application US/10311034
; Publication No. US20040023242A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: BANDMAN, Olga
; APPLICANT: BORSOWSKI, Mark L.
; APPLICANT: AO-YOUNG, Janice
; APPLICANT: LU, Yan
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: GREENWALD, Sara R.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: KEARNEY, Liam
; APPLICANT: BURFORD, Neil
; APPLICANT: NGUYEN, Dannel B.
; APPLICANT: TANG, Y. Tom
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: HE, Ann

```

; APPLICANT: THORNTON, Michael  
 ; APPLICANT: HAPALIA, April  
 ; APPLICANT: HARVIZU, Chandra S.  
 ; APPLICANT: GURURAJAN, Rajagopal  
 ; APPLICANT: LO, Terence P.  
 ; APPLICANT: KHAH, Farrah A.  
 ; APPLICANT: RECIPON, Shirley A.  
 ; APPLICANT: AZIMZAI, yaida  
 ; APPLICANT: POLICKY, Jennifer L.  
 ; APPLICANT: DING, Li  
 ; APPLICANT: GRETHER, Megan  
 ; APPLICANT: ELLIOTT, Vicki S.  
 ; APPLICANT: THANGAVELU, Kavitha  
 ; APPLICANT: BATRA, Sajeev  
 ; APPLICANT: ISON, Craig H.  
 ; TITLE OF INVENTION: HUMAN KINASES  
 ; FILE REFERENCE: PI-0125 PCT  
 ; CURRENT APPLICATION NUMBER: US/10/311.034  
 ; CURRENT FILING DATE: 2002-12-10  
 ; PRIORITY APPLICATION NUMBER: 60/212,073; 60/213,467; 60/215,651; 60/216,605; 60/218,372;  
 ; 60/228,056  
 ; PRIORITY FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-  
 ; 25  
 ; NUMBER OF SEQ ID NOS: 52  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 38  
 ; LENGTH: 3360  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. US20040023242A1 4841542CB1  
 ; US-10-311-034-38

Query Match	71.7%;	Score 1357.6;	DB 16;	Length 3360;
Best Local Similarity	83.7%;	Pred. No. 0;		
Matches 1569;	Conservative	0;	Mismatches 284;	Indels 21; Gaps 2;
QY	32	GCACAGGCTCCCTCGGCGCTCGCGCTCGAGGAGCGCCGCGCGCTGCGCGAGCGGC	91	
Db				
QY	98	GCGCGGCGCTTCGGCGCCCACTCCCTCGCGCAGAGCTAGCCCGCGCGCTGCGCGAAGGC	157	
Db				
QY	92	TCATCAAGTCGCGCTAAACCTCTGATGAAGAAGCAGGCGGTGAAGCGGCACCATCAAAAC	151	
Db	158	TGATCAAGTCGCGCCCAAGCCCTTAATGAAGAAGCAGGCGGTGAAGCGGCACCAACAAGC	217	
QY	152	ACAACCTGCGGCACCGGTACAGTTCTTGAGAGAGCTGGGCAAGGGCACCTACGGGAAGG	211	
Db	218	ACAACCTGCGGCACCGCTACAGTTCTTGAGAGACCTTGGGCAAAAGGCACCTACGGGAAGG	277	
QY	212	TGAAGAAGGCACGAGAGAGCTCGGGGCGTCTGGTGGCCATCAAGTCCATCAGAAAGACA	271	
Db	278	TGAAGAAGGCAGCGGAGAGCTCGGGGCGCTCGTGGCCATCAAGTCAATCCGGAAGACA	337	
QY	272	AAATCAAAGATGAGCAGGATCTGCTGCA CATACGAGGGAGATTGAGATCATCTCTTCAC	331	
Db	338	AAATCAAAGATGAGCAAGATCTGATGCACATACGAGGGAGATTGAGATCATCTCATCGC	397	
QY	332	TCACACACCCCA CATCATTTGCCATCCATGAAGTGTTTGAGATAGCAGCAAGATTGTGA	391	
Db	398	TCACACACCCCTCA CATCATTTGCCATCCATGAAGTGTTTGAGAACAGCAGCAAGATCGTGA	457	
QY	392	TTCTCATGTGAGTATGCCAGCGGGCGATCTGTATGATTACATCAGTGGCGCCACGGC	451	
Db	458	TCGTCTATGTGAGTATGCCAGCGGGGCGACCTTTATGACTACATCAGCGAGCGCGACGACG	517	
QY	452	TGAGTGAGCGGGACGCCAGGCAATTTCTTCGACAGATCGTGTCTGCCCTGCATCTAGTCGC	511	
Db	518	TCAGTGAGCGGAAGCTAGGCAATTTCTTCGCGAGATCGTCTCTGCCGTGCATTTGCC	577	
QY	512	ACCAGAACCGGATCGTTTCA CCGAGATCTCAAGCTGGAAAAACATCCTTCTAGATGCCAATG	571	
Db	578	ATCAGAACAGAGTTGTCCACCGAGATCTCAAGCTGGAGAACATCCTCTTGGGTGCCAATG	637	

```
Qy 1652 GTGAGGACAGCATCCTGTCTCCAGTCTTTGACCAATTGAGCTTCCTGAAAGTCTTC 1711
Db 1709 GCGAGGACAGCATCCTGTCTCTGAGTCTTTGACCAAGCTGGACTTCCTGAAAGGCTCC 1768
Qy 1712 CCGAAACCCCACTGAGGGGCTGTGTCTGTGGAACAACCTGAGGGGCTTGTGAGCAGCTC 1771
Db 1769 CAGAGCCCCCACTCGGGGCTGTGTCTGTGGAACAACCTCACGGGGCTTGTGAGGAGCCC 1828
Qy 1772 CCTCAGAAG-----GTCTGAAGCGATGGTGCAGGAATCCTTGGGGGATAGCT 1819
Db 1829 CCTCAGAGGGCCCTGGAGAGTCCCTGAGCGCTGGCGGAGGATCCTTTGGGGACAGCT 1888
Qy 1820 GCTTTTCTTGACAGACTGCCAAGAGGTGACTGTCAGGCTTACAGACAAGCCCTAGGAATCT 1879
Db 1889 GCTTTTCCCTGACAGACTGCCAAGAGGTGACAGCGACTTACCCGACAGGCACTGAGGGTCT 1948
Qy 1880 GCTCAAGCTCAGC 1893
Db 1949 GCTCAAGCTCACC 1962

RESULT 11
US-10-322-281-25
; Sequence 25, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ. ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3404
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-25

Query Match 71.7%; Score 1357; DB 17; Length 3404;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;

Qy 47 CCTCGCGCTGGCTCGGAGAGCGCCGCGCTGGCGGACGGCTCATCAAGTCGCTA 106
Db 162 CCATCTCCCTCGGCGCAGAGCTAGCCGCGCGCTGGCGGAGGGGCTGATCAAGTCGCCCCA 221
Qy 107 AACCTCTGATGAAGAAGCAGCGGCTGAAGCGGCAACCATCAAAACACAACTCGCGGCACC 166
Db 222 AGCCCTTANTGAAGACGAGCGGTGAAGCGGCACACACACAGCAAACTCGCGGCACC 281
Qy 167 GCTACGAGTTCCTGGAGACGCTGGGCAAGGCACTACGGGAAGGTGAAGAAGCAGCAG 226
Db 282 GCTACGAGTTCCTGGAGACCTGGGCAAGGCACTACGGGAAGGTGAAGAAGCAGCGGG 341
Qy 227 AGAGTCGGGCGTCTGGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGC 286
Db 342 AGAGTCGGGCGCTGGTGGCCATCAAGTCAATCCGGAAGGACAAATCAAGATGAGC 401
Qy 287 AGGATCTGCTGCATACGAGGAGGATTTGAGATCATGTCTTCACTCAACCAACCCCCACA 346
Db 402 AAGATCTGATGCAATACGAGGAGGATTTGAGATCATGTCTCACTCAACCCCTCACA 461
Qy 347 TCATTGGCATCCATGAAGTGTGAGAAATGACGAAGATGTGATGTGATGAGTATG 406
Db 462 TCATTGGCATCCATGAAGTGTGAGAAACAGCAGCAAGATCGTGTATGATGAGTATG 521
Qy 407 CCAGCCGAGGCGATCTGTATGATTAATCATCAGTACGCGGCACCGCTGAGTACGAGCGGAGC 466
Db 522 CCAGCCGAGGCGATCTTATGATTAATCATCAGTACGCGGCAGCAGCTCAGTACGAGCGGAG 581
Qy 467 CCAGGCAATTTCTCCGACAGATCGTGTCTGCGCTGCACTACTGCGCACACAGAACGGGATCG 526
```

```
Db 582 CTAGGCATTTCTTCCGCGAGATCGTCTCTGCGTGCATATTGCCATCAGAACAGAGTTG 641
Qy 527 TTCAACGAGATCTCAAGCTGGAAAAATCCTCTTAGATGCCAATGGAAAAATCAAGATTG 586
Db 642 TCACCCGAGATCTCAAGCTGGAAAAATCCTCTTGGATGCCAATGGAAAAATCAAGATTG 701
Qy 587 CTGACTTTGGCTCTCTCAACCTGTACCAAGGCAAGTTCTTCCAGACGTTCTGTGGGA 646
Db 702 CTGACTTTGGCTCTCTCAACCTGTACCAATCAAGGCAAGTTCTTCCAGACATTTCTGTGGGA 761
Qy 647 GCCTCTCTTACGGCTCGCTCGATAGTAGTCAACGGGAAAGCCCTATGTGGGCCGAGAGTGG 706
Db 762 GCCCTCTTATGCTCGCCAGAGATTGTCAATGGGAAGCCCTACACAGAGGCCGAGAGTGG 821
Qy 707 ACAGCTGGTCTCTGGGGGTTCTCTCTGTACATCTCTGTGTGATGCAACCATCCCTTTGACG 766
Db 822 ACAGCTGGTCTCTGGGGGTTCTCTCTGTACATCTCTGTGTGATGCAACCATCCCTTTGATG 881
Qy 767 GGCAGGATCATAAAAACACTGGTGAAGCAATCAGTAAACGGGGCTTACCGTGAGCGGCCCA 826
Db 882 GGCATGACCATAGATCCTTAGTGAACAGATCAGAACGGGGCTTACCGGAGGCCACTA 941
Qy 827 AGCGTCCGATGCTGTGGCTGATCGGTGGCTGTAAATGGTGAACCCCAACCGTGGGG 886
Db 942 AACCTCTGATGCTGTGGCTGATCGGTGGCTGTAAATGGTGAACCCCAACCGTGGGG 1001
Qy 887 CCACACTGGAGGATGATAGCCAGTCATGTGGGTCACTGGGTGTACACACCGAGTGC 946
Db 1002 CCACCTTGGAGATGTGGCAGTCACCTGGTGGGTCACTGGGTGTACACACCGAGTGG 1061
Qy 947 GGGAAACAGGAAGCCCTCGTGGGTGGGCAACCTAGTGTGTGCTTTGGCGGGCTTCCA 1006
Db 1062 GAGAGCAGGAGGCTCCGCATGAGGTGGGCAACCTGGCAGTGACTCTGCCGCGCTTCCA 1121
Qy 1007 TGGCGGACTGGTGTACGTCTCTCGCGCCCCCTCTCTGGAGAAATGGAGCCAAAGTGTGCA 1066
Db 1122 TGGCTGACTGGCTCCGCGCTTCTCTCGCGCCCCCTCTCTGGAGAAATGGGCGCAAGTGTGCA 1181
Qy 1067 GCTTCTTCAAGCAGCAGCTCGCGGAGGTGAGAGCACTGTACCTGGGCTGGAGCGGCAAC 1126
Db 1182 GCTTCTTCAAGCAGCAGTACCTGGTGGGGGAGAGCACCCCTGGGCTGGAGCGGCGAGC 1241
Qy 1127 ATTCTCTTAAGAAGTCCCGAAAGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGG 1186
Db 1242 ATTGCTCAAGAAATCCCGAAGAGAAATGACATGGCCAGTCTCTCCAAGTGAACAGG 1301
Qy 1187 CTGAGGATACCTCTTCTCGCCCTGGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATTCCTCA 1246
Db 1302 CTGATGACACTGCCCATCGCCCTGGCAAGAGCAACCTCAAGCTGCCAAAGGGCAATTCCTCA 1361
Qy 1247 AGAAAAAGTCTCTACCTCTGTCAGGGAGGTACAGGAGGACCTCAGGAATCAGACCGG 1306
Db 1362 AGAAGAAGGTGTGAGCTCTGCAAGAGGGGTACAGGAGGACCTCCGAGACTCAGCCCAA 1421
Qy 1307 TGCTGTATCTCAGAGGACGCTGTCCCTCTGTATCCCTGTCTCCCAAGAAAGGCAATCC 1366
Db 1422 TCCCTGGAGCCCAAGGACGGCTGCC-----CCCTGCTCCCCAGAGGGCAATTC 1472
Qy 1367 TTAAAGAGTCTCGACAGCGTGAATCTGTTTACTACTCTCTCCAGAGCCCCAGCGAGTCTG 1426
Db 1473 TCAAGAAGCCCCGACAGCGGAGTCTGGCTACTACTCTCTCCGAGCCCCAGTGAATCTG 1532
Qy 1427 GGGAACTCTTTAGAGCCAGTGTGTTTGTGAGTGGGAGCCCCGTGGAGGAGTCTC 1486
Db 1533 GGGAGCTCTTGGAGCGCAGCGAGCGTGTGTTGTGAGTGGGGATCCCAAGAGGAGGCTC 1592
Qy 1487 CACAGGCTTCAGGCGCTCTCTCCACCGCAAGGCGATTCTCAAACTCAATGGCAAGTTCT 1546
Db 1593 CGCAAGCTTCAGGCGCTCTCTCCATCGCAAGGCGATCTCAAACTCAATGGCAAGTTCT 1652
Qy 1547 CCGGCAAGCTTTAGAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCCT 1606
```



```
Db 1654 CGAAGCTTCAGGGCTGCTCTCCATCGCAAGGCATCCTCAAACTCAATGGCAAGTTCT 1713
Qy 1547 CCGGCACAGCCTTAGAAGGCACTACCTTAGCAGCTTTGGCTCCCTGGACCAACTGGCCT 1606
Db 1714 CCGAGACAGCCTTAGAGCTCGGGCCCCCAACCTTCGGCTCCCTGGATGAATCGGCC 1773
Qy 1607 CCTCCATCTCTGACGCGCGCCAGCCCTCAGGGGCTGTGAGTGAGACAGCATCC 1666
Db 1774 CACCTGGCCCTTGGCGCGGCGAGCGACCTCAGGGGCTGTGAGGAGACAGCATCC 1833
Qy 1667 TGTCTCCGAGTCTTTGACCAATTGAGCTTGCCTGAAAGCTTTCGCGAAACCCCACTGA 1726
Db 1834 TGTCTCTGAGTCTTTGACCAAGCTGGAATTTGCTGAACGGCTCCAGAGCCCCCACTGC 1893
Qy 1727 GGGGCTGTCTGTGTGACAACTGAGGGGCTTGGAGCAGCTCCCTCAGAAAG----- 1780
Db 1894 GGGGCTGTGTGTGTGTGACAACTCAGGGGCTTGGAGGAGCCCCCTCAGAGGGGCGCTG 1953
Qy 1781 -----GTCTGAAGCGATGTGGCAGGAATCCTTTGGGGGATAGCTGCTTTCTCTGACAG 1834
Db 1954 GAAGCTGCTTGAAGCGCTGGCGGCGAGGATCCTTTGGGGGACAGCTGCTTTCCCTGACAG 2013
Qy 1835 ACTGCCAAGGCTGACTGACGCTTACAGAACGCGCTTAGGAATCTGCTCAAAGCTCAGC 1893
Db 2014 ACTGCCAAGGCTGACAGCGACCTACCGACAGGCACTGAGGGTCTGCTCAAAGCTCACC 2072

RESULT 13
US-10-302-172-215
; Sequence 215, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Dmanac, Radjoje T.
; TITLE OF INVENTION: No. US20040053250A1el Arginine-rich Protein-like Nucleic Acids an
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803 1CNCp
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 215
; LENGTH: 2501
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(1989)
US-10-302-172-215

Query Match 71.6%; Score 1355.4; DB 16; Length 2501;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

Qy 47 CCTCGGCTTGGCTCGAGAGCGCCCGCGCTGGCGGAGCGGCTCATCAAGTCGCGTA 106
Db 140 CCACTCCCTCGGCGCGAGAGCTAGCCGCGCGCTGGCGGAGGGGCTGATCAAGTCGCGCA 199
Qy 107 AACCTCTGATGAAGAAGCAGCGGTGAAGCGGACCATCATCAAAACAAACCTCGGGCACC 166
Db 200 AGCCCTTAATGAAGAAGCAGCGGTGAAGCGGACCATCAAGCACAACCTCGGGCACC 259
Qy 167 GCTACGAGTCTCTGGAGACGCTGGGCAAGGACCTTAGCGGAAGGTTGAAGAGCAGAG 226
Db 260 GCTACGAGTCTCTGGAGACCTTGGGCAAGGACCTTAGCGGAAGGTTGAAGAGGCGCGG 319
```

```
Qy 227 AGAGCTCGGGCGCTCTGGTGGCCATCAAGTCCATCAGAAAGACAAAATCAAGATGAGC 286
Db 320 AGAGCTCGGGCGCTCTGGTGGCCATCAAGTCAATCCGGAGGACAAAATCAAGATGAGC 379
Qy 287 AGGATCTGCTGCACATACGAGGGAGATTGAGATCATGTCTTCACTCAACACACCCCAACA 346
Db 380 AAGATCTGATGCACATACGAGGGAGATTGAGATCATGTCTATCACTCAACACACCTCACA 439
Qy 347 TCATTGGCATCCATGAAGTCTTTGAGAAATAGCAGCAAGATTGTGATTGTCTAGAGATG 406
Db 440 TCATTGGCATCCATGAAGTCTTTGAGAAACAGCAGCAAGATCGTGATCGTCTAGGATG 499
Qy 407 CCAGCCGAGGCGATCTGTATGATTACATCAGTGGCGGCCACGGCTCAGTGAGCGGACG 466
Db 500 CCAGCCGAGGCGATCTTTATGACTACATCAGCGAGCGGAGCAGCTCAGTGAGCGGACG 559
Qy 467 CCAGGCAATTTCTTCGACAGATCGTGTCTCCCTGCACTACTGCCACAGAAACCGGATCG 526
Db 560 CTAGGCATTTCTTCGGCAGATCGTCTCTCGGTGCATATTGGCATCAGAAACAGAGTTG 619
Qy 527 TTCAACGAGATCTCAAGCTGGAAACATCTTCTTAGATGCCAATGGAAACATCAAGATTG 586
Db 620 TCCACCGAGATCTCAAGCTGGAAACATCTTCTTGGATGCCAATGGAAATATCAAGATTG 679
Qy 587 CTGACTTTGGCTCTCTCAACCTGTACCAAAAGCAAGTTCTCTCAGACGTTCTGTGGGA 646
Db 680 CTGACTTTGGCTCTCTCAACCTGTACCAACAGGCAAGTTCTTGCACACATCTGTGGGA 739
Qy 647 GCGCTCTCTAGGCTCGCTTGAGATAGTCAACGGGAAGCCCTATGTGGGCGGAGAGTGG 706
Db 740 GCGCTCTCTAGGCTCGCTTGAGATAGTCAACGGGAAGCCCTATCAGAGGCGGAGAGTGG 799
Qy 707 ACAGCTGGTCTCTGGGCGTTCTCTCTGTACATCTCTGTGTGATGGACCATGCGCTTTGACG 766
Db 800 ACAGCTGGTCTCTGGGCGTTCTCTCTGTACATCTCTGTGTGATGGACCATGCGCTTTGATG 859
Qy 767 GGCAAGGATCAATAAAGCACTGGTGAAGCAAACTAGTAACGGGCTTACCGTGAGCGGCCA 826
Db 860 GGCAAGGATCAATAAAGCACTGGTGAAGCAAACTAGTAACGGGCTTACCGTGAGCGGCCA 919
Qy 827 AGCGCTCCGATGCTGTGGCTGATCGGCTGCTTAAATGGTGAACCCCAACCGCTCGGG 886
Db 920 AACCTCTGATGCTGTGGCTGATCGGCTGCTTAAATGGTGAACCCCAACCGCTCGGG 979
Qy 887 CCACACTGGAGGATGTAGCCAGTCACTGGTGGGTCACTGGGTTACACACCGGAGTCG 946
Db 980 CCACCTCTGGAGGATGTAGCCAGTCACTGGTGGGTCACTGGGTTACACACCGGAGTCG 1039
Qy 947 GGGAAACAGGAAGCCCTGCGTGAGGCGGCAACCTAGTGTGACTTTGGCGGGGCTTCCA 1006
Db 1040 GAGAGCAGGAGGCTCGCATGAGGGTGGGCAACCTTGGCAGTGACTCTGCCGCGCTTCCA 1099
Qy 1007 TGGCGGACTGCTTACGCTCGCTCGCGCCCTCTCTGGAGAAATGGAGCAAGGTTGCA 1066
Db 1100 TGGCTGACTGGCTCGGCGTTCTCTCGCGCCCTCTCTGGAGAAATGGGCGCAAGGTTGCA 1159
Qy 1067 GCTTCTTCAAGCAGCAGCTCGCGGAGGTGGAAGCACTGTACTGGGCTCGAGGCGGCAAC 1126
Db 1160 GCTTCTTCAAGCAGCAGCTCGCGGAGGTGGAAGCACTGTACTGGGCTCGAGGCGGAGC 1219
Qy 1127 ATTCTCTTAAGAAAGTCCGAAAGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGG 1186
Db 1220 ATTCTCTTAAGAAAGTCCGAAAGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGG 1279
Qy 1187 CTGAGGATACCTTCTCGGCTTGGCAAGAGCAGCTTAAAGCTTTCGAAAGGCAATTTCTCA 1246
Db 1280 CTGATGACATGCGCCCTCGGCTTGGCAAGAGCACTTCAAGCTGCGCAAGGCAATTTCTCA 1339
Qy 1247 AGAAAGAGTCTCTACTCTGTAGGGGAGGTACAGGAGGACCTCAGGAACTCAGAGCGG 1306
Db 1340 AGAAGAGAGTCTCAGCCTCTGCAAGAGGGGTACAGGAGGACCTCAGGAGCTCAGGCCAA 1399
```

1307	QY	TGCTTGATCTACGAGCGACCTGTCCCTGCTGTATCTCTGCTCCCAAGAAAGGCATCC	1366
1400	Db	TCCCTCGAGCCACGAGCGAGGCTGCC-----CCCTGCTCCCAAGAGGCGATTC	1450
1367	QY	TTAAGAAGTCTCGACAGCGTGAATCTGGTTACTTACTCTCTCCAGAGCCCGAGTCTG	1426
1451	Db	TCAAGAAGCCCCGACAGCGGAGTCTGGCTACTACTCTCTCCGAGCCCGAGTGAATCTG	1510
1427	QY	GGGAACTCTTAGACGCCAGTGATGTGTTCTGTAGTGGGAGCCCGTGGAGCAGAAGTCTC	1486
1511	Db	GGGAGCTCTTTGGACGACGAGCGAGCTGTTTGTGAGTGGGGATCCCAAGGACGAGAGCCTC	1570
1487	QY	CACAGGCTTCAGGGCTCCTCTCCACGCGAAGGCGATTTCTCAAACCTCAATGGCAAGTTCT	1546
1571	Db	CGCAAGCTTCAGGGCTGCTCCTTCATCGCAAGAGCATCTCTCAAACCTCAATGGCAAGTTCT	1630
1547	QY	CCGCGACAGCTTTAGAAGGCACTATCCCTTAGCACCTTTTGCTCCCTGGACCAACTGGCCT	1606
1631	Db	CCCAGACAGCTTTGGAGCTCGCGGCCCCACCACTTTGCGCTCCTTGATGAACTCGCCC	1690
1607	QY	CCCTCCCATCTGCAGAGCCCGCCAGCGCCCTTCAGGGGCTGTGAGTGAGGACAGCATCC	1666
1691	Db	CACCTGCCCTCTGGCCCGGGCCAGCCGACCTCAGGGGCTGTGAGCGAGGACAGCATCC	1750
1667	QY	TGTCCTCCGAGTCCCTTTGACCAATTGCACTTGCTCGTGAACGCTTCTCCGAAACCCCACTGA	1726
1751	Db	TGTCTCTGAGTCCCTTTGACCGAGCTGACTTGCTCGAACGGCTCCACAGAGCCCCCACTGC	1810
1727	QY	GGGGCTGTGTGTGTGGACAACTGAGGGGGCTTTAGCAGCCCTCCCTCAGAG-----	1780
1811	Db	GGGGCTGTGTGTGTGGACAACTCA CGGGGCTTGAGGAGCCCCCTCAGAGGGGCCCTG	1870
1781	QY	-----GTCTGAAGCGATGGTGGCAGGAATCCTTTGGGGGATAGTGTCTTTCTCTGACAG	1834
1871	Db	GAAGCTGCCTGAGCGCTGGCGGACAGATCTCTTTGGGGGACAGTGTCTTTCCCTGACAG	1930
1835	QY	ACTGCCAAGAGGTGACTGCAGCCTACAGACAAGCCCTAGGAATCTGCTCAAAGTCTAGC	1893
1931	Db	ACTSCCAGGAGGTGACACGACCTTACCGACAGGCACCTCAGGGTCTGCTCAAAGCTCACC	1989

```

RESULT 14
US-10-370-715B-639
; Sequence 639, Application US/10370715B
; Publication No. US20040258678A1
; GENERAL INFORMATION:
;   Patin Docket Preview
; APPLICANT: BODARY, SARAH C.
; APPLICANT: CLARK, HILARY
; APPLICANT: BRISDELL, HUNTE
; APPLICANT: JACKMAN, JANET
; APPLICANT: SCHOENFELD, JILL R.
; APPLICANT: WILLIAMS, P. MICKEY
; APPLICANT: WOOD, WILLIAM I.
; APPLICANT: WU, THOMAS D.
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune
; TITLE OF INVENTION: Related Diseases
; FILE REFERENCE: P1948R1-US
; CURRENT APPLICATION NUMBER: US/10/370,715B
; CURRENT FILING DATE: 2003-02-21
; NUMBER OF SEQ ID NOS: 742
; SEQ ID NO 639
; LENGTH: 3443
; TYPE: DNA
; ORGANISM: Homo sapien
; US-10-370-715B-639

```

Db	162	CCACTCCCTCGGCGCGAGCTAGCCCGCGCTGGCGGAAGGCTGATCAAGTCGCCCA	221
Qy	107	AACCTCTGATGAAGACAGCGGCTGAAGCGGCACATCACAAAACAAACCTTGGCGACCC	166
Db	222	AGCCCCAATGAAGAGCAGCGGCTGAAGCGGCACACACAGCACAACTTGGCGACCC	281
Qy	167	GCTACAGATTCTGGAGACGCTGGCGCAGGCGACCTACGGGAAGGTGAAGAAGGCACGAG	226
Db	282	GCTACAGATTCTGGAGACCCCTGGGCAAGGCACTTACGGGAAGGTGAAGAAGGCACG	341
Qy	227	AGAGCTCGGGGCGCTCTGTGTGGCCATCAAGTCCATCAGGAAGACAAAAATCAAAGATGAGC	286
Db	342	AGAGCTCGGGGCGCTGTGTGGCCATCAAGTCCATCGGAAGACAAAAATCAAAGATGAGC	401
Qy	287	AGGATCTGTGCACATACCGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAACA	346
Db	402	AAGATCTGATGCATATACGAGGGGAGATTGAGATCATGTCTTCACTCAACCAACCCCTACA	461
Qy	347	TCATTGCCATCCATGAAGTGTGGAAATAGCAGCAAGATTGTGATTGTTCATGGAGTATG	406
Db	462	TCATTGCCATCCATGAAGTGTGGAAATAGCAGCAAGATTGTGATTGTTCATGGAGTATG	521
Qy	407	CCAGCCGAGCGCATCTGTATGATTACATCAGTGAGCGGCCACGGCTAGTGAGCGGGACG	466
Db	522	CCAGCCGAGCGCATCTTATGACTACATCAGCGAGCGGAGCAGCTCAGTGAGCGCGAAG	581
Qy	467	CCAGGCATTCTTCGACAGATCGTCTGCGCTGCACCTACTGCCACCAAGACGGGATCG	526
Db	582	CTAGGCATTCTTCCGCGAGATCGTCTGCGCTGCACCTATTGGCCATCAGACAGAGTTG	641
Qy	527	TTCAACCGAGATCTCAAGCTGGAAACATCTTCTAGATGCCAATGGAAACATCAAGATTG	586
Db	642	TCCACCGAGATCTCAAGCTGGAGAACATCTCTTTGGATGCCAATGGAAATCAAGATTG	701
Qy	587	CTGACTTTGGCCTCTCGAACCTGTACCAAGAGCAAGTTCTCTCAGAGCTTCTGTGGGA	646
Db	702	CTGACTTTGGCCTCTCGAACCTGTACCATCAAGGCAAGTTCTCTGAGACATCTGTGGGA	761
Qy	647	GCCCTCTACGCGCTCGCCTAGATAGTCAACGGGAAGCCCTATGTGGGCCACAGAGTGG	706
Db	762	GCCCCCTCATGCTCGCCAGAGATTGTCAAATGGGAAGCCCTACACAGGCCACAGAGTGG	821
Qy	707	ACAGCTGGTCTCTGGGGGTCTCTGTACATCTGTGTGCATGSCACATGCCCTTTTACG	766
Db	822	ACAGCTGGTCTCTGGGGTCTCTCTCTACATCTGTGTGCATGSCACATGCCCTTTTATG	881
Qy	767	GGCAGGATATAAAACACTGCTGAAGCAATCAGTAACGGGGCTTACCGGTGAGCCGCCCA	826
Db	882	GGCATGACCAATAGATCTTAGTGAACAGATCAGCAACGGGGCTTACCGGGACCACTA	941
Qy	827	AGCCGTCGATGCTGTGGCTGATCCGGTGGCTGTGTTAATGTTGAACCCCAACCCGTCGG	886
Db	942	AACCCCTCTGATGCTGTGGCTGATCCGGTGGCTGTGTTGTTGTTGAACCCCAACCCGCG	1001
Qy	887	CCACATCTGGAGATGACGAGTCATTGTTGGGTCAACTGGGGTTACACCAACCGAGTCG	946
Db	1002	CCACCCCTGGAGGATGTGGCCAGTCACTGTGTGGGTCAACTGGGGGTACGCCCAACCGAGT	1061
Qy	947	GGGAACAGGAAGCCCTCGCTGAGGGTGGGCACCCCTAGTGTGATTTTGGCCGGCCCTCCA	1006
Db	1062	GAGAGCAGGAGGCTCCGCATGAGGGTGGGCACCTCTGGCAGTGACTCTGTCCCGCGCTCCA	1121
Qy	1007	TGGCGGACTGTTAGTTCGCTCTCGCGCCCCCTCTGGAGATGGAGCAAGGTGTGCA	1066
Db	1122	TGGCTGATGCTCTCGCGCTTCTCTCGCGCCCCCTCTCTGGAGATGGGGCCAAAGTGTGCA	1181
Qy	1067	GTCTTCTTCAAGCAGCAGTGTCCGGGAGGTGGAAAGCACTGTACCTTGGGCTTGGAGCGCAAC	1126
Db	1182	GTCTTCTTCAAGCAGCAGTGTCTCTGGTGGGGAGGACCAACCCCTTGGCTTGGAGCGCGAG	1241
Qy	1127	ATTCTCTTAAAGATCCCGAAGAGAGATGACATGGCTCAAAATCTGCAAGGTGACCCCG	1186
Db	1242	ATTGGCTCAAGAGTCCCGAAGAGATGACATGGCCCCAGTCTCTTCAATGACACGG	1301



Qy 1187 CTGAGGATACCTCTTCTCGCCCTGGCAAGACGAGCCTTAAGCTTCCGAAAGGCATTCTCA 1246  
Db 1302 CTGATGACACTGCCATCGCCCTGGCAAGACCACTCAAGCTGCCAAGGGCAATTCTCA 1361  
Qy 1247 AGAAAAGTCTCTTACCTCTGTCAGGGAGGTACAGGAGGACCTTCAGAACTCAGACCGG 1306  
Db 1362 AGAAGAAGGTGTACGCTCTGTCAGAAAGGGGTACAGGAGGACCTCCGAGCTCAGGCCAA 1421  
Qy 1307 TGCCTGATACCTCAGGCGAGCTGTCCCTGCTGTATCCCTGCTCCCAAGGAAGGCATCC 1366  
Db 1422 TCCCTGGAGCCCAAGGCGAGCTGCC-----CGCTGCTCCCAAGAGGGCATTC 1472  
Qy 1367 TTAAGAAGTCTCGACAGCGGTGAATCTGTGTTACTACTCTCTCCAGAGCCAGCGAGTCTG 1426  
Db 1473 TCAAGAAGCCCGACAGCGGAGTCTGGCTACTACTCTCTCCGAGCCCAAGTGAATCTG 1532  
Qy 1427 GGGAACTCTTAGACCGCAGTGAATGTGTTGTGAGTGGGACCCCGTGGAGCAGAAGTCTC 1486  
Db 1533 GGGAGCTCTTGGACGCGAGCGACGTGTTGTGAGTGGGATCCCAAGGAGCAGAAGCCTC 1592  
Qy 1487 CACAGGCTTCAAGGCTCTCTCTCCACCGCAGGCGATCTCAAACTCAATGGCAAGTCT 1546  
Db 1593 CGCAAGCTTCAAGGCTGTCTCTCATCGCAAGGCACTCTCAAACTCAATGGCAAGTCT 1652  
Qy 1547 CCCGCACAGCCTTAGAAGGCATACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCCT 1606  
Db 1653 CCCAGACAGCCTTGAGCTCGCGGCCCCCAACCTCTGGCTCCCTGGATGAATCTCGCC 1712  
Qy 1607 CCTCCATCTCGACGCGCGGCCAGCGCCCTCAGGGGCTGTGAGTGAGACAGCATCC 1666  
Db 1713 CACCTCGCCCTCGCGGCCAGCGCACTCTCAGGGGCTGTGAGCGAGCAGCATCC 1772  
Qy 1667 TGTCTCCGAGTCTTGAACCAATTTGACTGTGCTGACGCTCTTCCGGAACCCCACTGA 1726  
Db 1773 TGTCTCTGAGTCTCTTGACCACTGTGACTTGTGCTGAACTGCTGAACTGCTGCTGCTGCTG 1832  
Qy 1727 GGGGCTGTGTCTGTGTGCAACCTTGAGGGGGCTTGAGCAGCCTCTCTCAGAAAG----- 1780  
Db 1833 GGGGCTGTGTCTGTGTGCAACCTTGAGGGGGCTTGAGGAGCCCTCTCAGAGGGCCCTG 1892  
Qy 1781 -----GTCTGAAGCGATGTGGCAGGAATCCTTTGGGGGATPAGTGTCTTTCTGACAG 1834  
Db 1893 GAAGCTGCTGAGGGCTGGCGGAGGATCTCTTTGGGGGACAGTGTCTTTCCCTGACAG 1952  
Qy 1835 ACTGCCAAGGTCAGTGCAGCCTACAGACAGCCCTAGGATCTGCTCAAGCTCAGC 1893  
Db 1953 ACTGCCAAGGTCAGCAGCAGCTACCGACAGGCACTGAGGGTCTGCTCAAGCTCACC 2011

RESULT 15

US-10-343-514-13  
; Sequence 13, Application US/10343514  
; Publication No. US20040132025A1  
; GENERAL INFORMATION:  
; APPLICANT: DRUCKER, Daniel J.  
; APPLICANT: ROSEN, Cheryl F.  
; APPLICANT: LEBEVRE, Diana L.  
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK  
; FILE REFERENCE: DPA-DRUC2/PCT  
; CURRENT APPLICATION NUMBER: US/10/343,514  
; PRIOR FILING DATE: 2003-01-31  
; PRIOR FILING DATE: 2003-01-31  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 109  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 13

; LENGTH: 1186  
; TYPE: DNA  
; ORGANISM: RAT  
US-10-343-514-13

Query Match 56.6%; Score 1071.4; DB 17; Length 1186;  
Best Local Similarity 94.0%; Pred. No. 0;  
Matches 1114; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 70 GCCCGCGCTGGCGGACGGGCTCATCAAGTCGGCTAAACCTCTGTATGAAGAAGCAGCGG 129  
Db 1 GCCCGCGCTGGCGGACGGGCTCATCAAGTCGGCTAAACCTCTGTATGAAGAAGCAGCGG 60  
Qy 130 GTGAAGCGGACCATCAAAACAACCTCTGGGACCGCTACGAGTTCCTGGAGACGCTG 189  
Db 61 GTGAAGCGGACCATCAAAACAACCTCTGGGACCGCTACGAGTTCCTGGAGACGCTG 120  
Qy 190 GGCAAGGGCACCTACGGGAAGGTGAAGAAGCAAGAGAGCTCGGGCGCTCTGGTGGCC 249  
Db 121 GGCAAGGGCACCTACGGGAAGGTGAAGAAGCAAGAGAGCTCGGGACGCTCTGGTGGCC 180  
Qy 250 ATCAAGTTCATCAGGAAGCAAAATCAAAAGATGAGCAGGATCTGTGCAATACGAGG 309  
Db 181 ATCAAGTTCATCAGGAAGCAAAATCAAAAGATGAGCAGGATCTGTGTCATTAAGGAG 240  
Qy 310 GAGATTGAGATCATGTCTTCACTCAACCCACATCATTTGCCATCCATGAAGTGTCT 369  
Db 241 GAGATCGAGATCATGTCTTCACTCAACCCACATCATTTGCCATCCATGAAGTGTCT 300  
Qy 370 GAGAATAGCAGCAAGATTGTGATTGTCATGAGGATATGCCAGCGAGGCGATCTGTATGAT 429  
Db 301 GAGAATAGCAGCAAGATTGTGATTGTCATGAGGATATGCCAGCGAGGCGATCTGTATGAT 360  
Qy 430 TACATCAGTACGCGGCAACCGCTGAGTGAGCGGGAAGCGGAGGATTTCTTCCGACAGATC 489  
Db 361 TACATCAGTACGCGGCAACCGCTGAGTGAGCGGGAAGCGGAGGATTTCTTCCGACAGATC 420  
Qy 490 GTGTCTGCCCTGCACTACTGCCACCAAGCGGATCGTTTCAACGAGATCTCAAGCTCGAA 549  
Db 421 GTGTGCCCTGCACTACTGCCACCAAGCGGATCGTTTCAACGAGATCTCAAGCTCGAG 480  
Qy 550 AACATCTCTCTAGATGCAATGGAACATCAAGATTGCTGACTTTGGGCTCTCCAACTCG 609  
Db 481 AACATCTCTCTAGATGCAATGGAACATCAAGATTGCTGACTTTGGGCTCTCCAACTCG 540  
Qy 610 TACCAAAAGCGAAGTTCCTCAGACGTTCTGTGGAGCCCTCTCTACGCTCGCCTGAG 669  
Db 541 TATCAAAAGCGAAGTTCCTCAGACGTTCTGTGGAGCCCTCTCTATGGCTCTCACTGAG 600  
Qy 670 ATAGTCAAGCGGAAGCCCTATGTGGGCCAGAGGTGGACAGCTGGTCTCTGGGCGTTCTC 729  
Db 601 ATCGTCAAGCGGAAGCCCTATGTGGGCCAGAGGTGGACAGCTGGTCTCTGGGCGTTCTC 660  
Qy 730 CTGTACATCTCTGTGATGSCACCATGCGCTTTTGAAGCGGAGGATCAATAAACAACCTGGT 789  
Db 661 CTGTACATCTCTGTGATGSCACCATGCGCTTTTGAAGCGGAGGATCAATAAACAACCTGGT 720  
Qy 790 AAGCAATCAGTAAACGGGCTTACCGTACCGCCCGCCAGCGCTCCGATGCTGTGGCGCTG 849  
Db 721 AAGCAATCAGTAAACGGGCTTACCGTACCGCCCGCCAGCGCTCCGATGCTGTGGCGCTG 780  
Qy 850 ATCCGCTGGCTGTAAATGGTGAACCCCAACCGCTCGGSCACACTGGAGGATGTAGCCAGT 909  
Db 781 ATCCGCTGGCTGTAAATGGTGAACCCCAACCGCTCGGSCACACTGGAGGATGTAGCCAGT 840  
Qy 910 CATTTGGTGGTCAACTGGGTTTACCAACCGGAGTGGGGAACAGGAAGCCCTCGCTGAG 969  
Db 841 CATTTGGTGGTCAACTGGGTTTACCAACCGGAGTGGGGAACAGGAAGCTCTGCGGAGAG 900  
Qy 970 GGTGGGCAACCTAGTGGTGTGCTTTGGCGGGGCTCCATGGCGGAGCTGGTTACGTCGCTCC 1029  
Db 901 GGTGGGCAACCTAGCGGTGATCTGTGGCGGGGCTCTATGGCGGAGCTGTTTACGTCGCTCC 960



Qy	1030	TCGGGCCCCCTCTCGAGAAATGGAGCAAGGTGTGCAGCTTCTTCAAGAGCAGCTGCCG	1089
Db	961	TCGGGCCCCCTCTCGAGAAATGGAGCAAGGTGTGCAGCTTCTTCAAGAGCAGCTGCCG	1020
Qy	1090	GGAGGTGGAGCACTGTACCTTGGGCTGGAGCGGCACATTCTCTTAAGAAGTCCCGAAG	1149
Db	1021	GGAGTGGAGACACGGGACCGGGCTGGAGCGGCACATTCTCTTAAGAAGTCCCGCAG	1080
Qy	1150	GAGAAATGACATGGCTCAAAATCTCGAAGGTGACCCGGCTGAGGATACCTTCTTCGCCCT	1209
Db	1081	GAGAAATGACATGGCTCAGACTCTGCAGAAATGACCCAGTGAAGATACTTCTCTCGCCCT	1140
Qy	1210	GGCAAGAGCAGCCTTAAGCTTCCGAAAGGCATTTCTCAAGAAAAG	1254
Db	1141	GGCAAGAACAGCCTCAAGCTTCCGAAAGGTATCTCAAGAAAAG	1185

Search completed: January 25, 2005, 07:55:10  
Job time : 996.334 secs



QY 361 GAAGTGTGTTGAGAAATAGCAGCAAGATTGTTGATTTGTCATGGAGTATGCCAGCCGAGGGCAT 420  
DB 483 GAAGTGTGTTGAGAAATAGCAGCAAGATTGTTGATTTGTCATGGAGTATGCCAGCCGAGGGCAT 542  
QY 421 CTGTATGATTACATCAGTACGAGCGGCCACCGGTGAGTACGCGGGACCGCCAGGCAATTTCTTC 480  
DB 543 CTGTATGATTACATCAGTACGAGCGGCCACCGGTGAGTACGCGGGACCGCCAGGCAATTTCTTC 602  
QY 481 CGACAGATCGTGTCTGCCCTGCATCTGCGCACACAGAAACGGGATCGTTTACCGAGATCTC 540  
DB 603 CGACAGATCGTGTCTGCCCTGCATCTGCGCACACAGAAACGGGATCGTTTACCGAGATCTC 662  
QY 541 AAGCTGGAACATCTCTTAGATGCGCAATGGAACATCAAGATTGCTGACTTTGGGCTC 600  
DB 663 AAGCTGGAACATCTCTTAGATGCGCAATGGAACATCAAGATTGCTGACTTTGGGCTC 722  
QY 601 TCCAACTGTACCAAAAGGCAAGTTCTCTCAGACGTTCTGTGGGAGCCCTCTCTAGGCC 660  
DB 723 TCCAACTGTACCAAAAGGCAAGTTCTCTCAGACGTTCTGTGGGAGCCCTCTCTAGGCC 782  
QY 661 TCGCCTGAGTAGTCAACGGGAAGCCCTATGTGGGCCACAGAGTGAGCAGCTGTCTCTG 720  
DB 783 TCGCCTGAGTAGTCAACGGGAAGCCCTATGTGGGCCACAGAGTGAGCAGCTGTCTCTG 842  
QY 721 GGCCTTCTCCTGTACATCTCTGTGATGCGCACCATGCTTTGACGGGCGAGATCATAAA 780  
DB 843 GGCCTTCTCCTGTACATCTCTGTGATGCGCACCATGCTTTGACGGGCGAGATCATAAA 902  
QY 781 ACCTGTGTGAAGCAAAATCAGTAACGGGCTTTACCGTGAGCGCCCAAGCCCTCGATGCC 840  
DB 903 ACCTGTGTGAAGCAAAATCAGTAACGGGCTTTACCGTGAGCGCCCAAGCCCTCGATGCC 962  
QY 841 TGTGGCTGTATCCCGTGGCTGTTAATGTGTAACCCCAACCGTCCGGGCCACACTGGAGAT 900  
DB 963 TGTGGCTGTATCCCGTGGCTGTTAATGTGTAACCCCAACCGTCCGGGCCACACTGGAGAT 1022  
QY 901 GTAGCCAGTCAATGTGGGTCAAATGGGGTTTACACACCGGAGTGGGGACAGGAAGCC 960  
DB 1023 GTAGCCAGTCAATGTGGGTCAAATGGGGTTTACACACCGGAGTGGGGACAGGAAGCC 1082  
QY 961 CTGGCTGAGGTGGGACCCCTAGTGGTGAATTTGGCCGGGCTCCATGGCGGACTGGTTA 1020  
DB 1083 CTGGCTGAGGTGGGACCCCTAGTGGTGAATTTGGCCGGGCTCCATGGCGGACTGGTTA 1142  
QY 1021 CGTGGCTCTCCGCGCCCTCTCGAGNATGGAGCCAGGTGTGACGTTCTTCAAGCAG 1080  
DB 1143 CGTGGCTCTCCGCGCCCTCTCGAGNATGGAGCCAGGTGTGACGTTCTTCAAGCAG 1202  
QY 1081 CAGTGGCGGAGGTGGAAGCACTGTACTGGGCTGGAGCGGCAACATTTCTTTAAGAAG 1140  
DB 1203 CAGTGGCGGAGGTGGAAGCACTGTACTGGGCTGGAGCGGCAACATTTCTTTAAGAAG 1262  
QY 1141 TCCGGAAGGAGATGATGGCTCAAATCTGCAAGGTGACCCGGCTGAGGATACCTCT 1200  
DB 1263 TCCGGAAGGAGATGATGGCTCAAATCTGCAAGGTGACCCGGCTGAGGATACCTCT 1322  
QY 1201 TCTCGCCCTCGCAGAGCAGCTTAAGTTCGGAAGGCAATTTCAAGAAAGTCTCT 1260  
DB 1323 TCTCGCCCTCGCAGAGCAGCTTAAGTTCGGAAGGCAATTTCAAGAAAGTCTCT 1382  
QY 1261 ACCTGTGAGGGAGGTACAGGAGGACCCCTCAGAACTCAGACCGGTGCTGTACTTCCA 1320  
DB 1383 ACCTGTGAGGGAGGTACAGGAGGACCCCTCAGAACTCAGACCGGTGCTGTACTTCCA 1442  
QY 1321 GGGAGCCTGTCCCTGTGTATCTCTGCTCCAGGAAGGCAATCTTAAGAGTCTCGA 1380  
DB 1443 GGGAGCCTGTCCCTGTGTATCTCTGCTCCAGGAAGGCAATCTTAAGAGTCTCGA 1502  
QY 1381 CAGCGTGAATCTGTTACTACTCTCTCAGAGCCAGCGAGTCTGGGAACTCTTAGAC 1440  
DB 1503 CAGCGTGAATCTGTTACTACTCTCTCAGAGCCAGCGAGTCTGGGAACTCTTAGAC 1562  
QY 1441 GCCAGTGAATGTGTTGTAGTGGGGACCCCGTGGAGCAGAAAGTCTCCACAGGCTTCAGGG 1500

DB 1563 GCCAGTGAATGTGTTGTAGTGGGGACCCCGTGGAGCAGAAAGTCTCCACAGGCTTCAGGG 1622  
QY 1501 CTCTCTCCACCGCAAGGGCAATTTCAAACCTCAATGGCAAGTTCTCCCGCACAGCCTTA 1560  
DB 1623 CTCTCTCCACCGCAAGGGCAATTTCAAACCTCAATGGCAAGTTCTCCCGCACAGCCTTA 1682  
QY 1561 GAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCCTCTCCCATCTCTGCA 1620  
DB 1683 GAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCCTCTCCCATCTCTGCA 1742  
QY 1621 GCCCGGCCAGCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTCTCTCCGAGTCC 1680  
DB 1743 GCCCGGCCAGCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTCTCTCCGAGTCC 1802  
QY 1681 TTTGACCAATTTGACCTTGGCTGAACGCTTTCCCGAAACCCCACTGAGGGGCTGTGTGCT 1740  
DB 1803 TTTGACCAATTTGACCTTGGCTGAACGCTTTCCCGAAACCCCACTGAGGGGCTGTGTGCT 1862  
QY 1741 GTGGACAACTGAGGGGCTTTGAGCAGCCTCTCTCAGAAAGTCTGAAGCGATGGTGGCAG 1800  
DB 1863 GTGGACAACTGAGGGGCTTTGAGCAGCCTCTCTCAGAAAGTCTGAAGCGATGGTGGCAG 1922  
QY 1801 GAATCTTTGGGGATAGCTGCTTTTCTGACAGACTGCGCAAGAGTGTGACGAGCTAC 1860  
DB 1923 GAATCTTTGGGGATAGCTGCTTTTCTGACAGACTGCGCAAGAGTGTGACGAGCTAC 1982  
QY 1861 AGACAAGCCCTAGGAATCTGCTCAAAAGCTCAGC 1893  
DB 1983 AGACAAGCCCTAGGAATCTGCTCAAAAGCTCAGC 2015

RESULT 2  
US-10-355-975A-4  
; Sequence 4, Application US/10355975A  
; Patent No. 6759223  
; GENERAL INFORMATION:  
; APPLICANT: Immunex Corporation  
; APPLICANT: Bird, Timothy A.  
; APPLICANT: Virca, G. Duke  
; APPLICANT: Martin, Unja  
; APPLICANT: Anderson, Dirk M.  
; TITLE OF INVENTION: CALCIUM/CALMODULIN-DEPENDENT KINASE  
; FILE REFERENCE: 2923-B  
; CURRENT APPLICATION NUMBER: US/10/355,975A  
; CURRENT FILING DATE: 2003-01-30  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 2902  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-355-975A-4

Query Match 100.0%; Score 1893; DB 4; Length 2902;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1893; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGTCCGTTGGCCTTACTTCCAGCGCCCGAGCAGGCTCCCTCGGCTCCGCGCTGGCC 60  
DB 123 ATGGAGTCCGTTGGCCTTACTTCCAGCGCCCGAGCAGGCTCCCTCGGCTCCGCGCTGGCC 182  
QY 61 TCGGAGAGCGCCCGGCGCTTGGGAGCGGGCTCATCAAGTCGCTTAACTCTGTATGAAG 120  
DB 183 TCGGAGAGCGCCCGGCGCTTGGGAGCGGGCTCATCAAGTCGCTTAACTCTGTATGAAG 242  
QY 121 AAGCAGCGGTGAAGCGGCAACATCAAAACAACTCGCGCACCGCTACGAGTTCCTG 180  
DB 243 AAGCAGCGGTGAAGCGGCAACATCAAAACAACTCGCGCACCGCTACGAGTTCCTG 302  
QY 181 GAGACGCTGGGCAAGGGCAACCTACCGGAAGGTGAAGAGGCAAGAGAGCTCGGGCGGT 240  
DB 303 GAGACGCTGGGCAAGGGCAACCTACCGGAAGGTGAAGAGGCAAGAGAGCTCGGGCGGT 362

QY 241 CTGTGGCCATCAAGTCCATCAGGAAGACAAATCAAAGATGAGCAGGATCTGTGCAC 300  
DB 363 CTGTGGCCATCAAGTCCATCAGGAAGACAAATCAAAGATGAGCAGGATCTGTGCAC 422  
QY 301 ATACGGAGGAGATGAGATCATGTCTTCACTCAACCAACCCCAACATATTCGCATCAT 360  
DB 423 ATACGGAGGAGATGAGATCATGTCTTCACTCAACCAACCCCAACATATTCGCATCAT 482  
QY 361 GAAGTGTGAGAAATAGCAGCAAGATGTGATGTCATGAGATGACAGCCGAGGCGAT 420  
DB 483 GAAGTGTGAGAAATAGCAGCAAGATGTGATGTCATGAGATGACAGCCGAGGCGAT 542  
QY 421 CTGTATGATTAACATCAGTGAAGCCGCGCTGAGTGAAGCGGAGCGGAGGCAATTTCTTC 480  
DB 543 CTGTATGATTAACATCAGTGAAGCCGCGCTGAGTGAAGCGGAGCGGAGGCAATTTCTTC 602  
QY 481 CGACAGATCGTGTCTGCGCTGCATCTACTGCCACAGAACGGGATCGTTACCGAGATCTC 540  
DB 603 CGACAGATCGTGTCTGCGCTGCATCTACTGCCACAGAACGGGATCGTTACCGAGATCTC 662  
QY 541 AAGCTGGAACATCTCTTAGATGCCAATGGAACATCAAGATGCTGACTTTGGGCTC 600  
DB 663 AAGCTGGAACATCTCTTAGATGCCAATGGAACATCAAGATGCTGACTTTGGGCTC 722  
QY 601 TCCAACCTGTACCAACAAAGGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTCTCTAGGCC 660  
DB 723 TCCAACCTGTACCAACAAAGGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTCTCTAGGCC 782  
QY 661 TCGCTGAGATGATCAACGGGAAGCCCTATGTGGGCCAGAGTGGACAGTGTCTCTG 720  
DB 783 TCGCTGAGATGATCAACGGGAAGCCCTATGTGGGCCAGAGTGGACAGTGTCTCTG 842  
QY 721 GGCCTTCTCTGTACATCTCTGTCATGGCAGCACCATGCCCTTTGACGGGAGGATCAATAA 780  
DB 843 GGCCTTCTCTGTACATCTCTGTCATGGCAGCACCATGCCCTTTGACGGGAGGATCAATAA 902  
QY 781 AACTGTGTGAAGCAAAATCAGTAACGGGGCTTACCGTGAGCGCCCAAGCCGCTCCGATGCC 840  
DB 903 AACTGTGTGAAGCAAAATCAGTAACGGGGCTTACCGTGAGCGCCCAAGCCGCTCCGATGCC 962  
QY 841 TGTGGCTGTATCCGGTGGCTGTTAATGGTGAAACCCCAACCGCTCGGGCCACACTGGAGAT 900  
DB 963 TGTGGCTGTATCCGGTGGCTGTTAATGGTGAAACCCCAACCGCTCGGGCCACACTGGAGAT 1022  
QY 901 GTAGCCAGTCAATGGTGGCTCAACTGGGTTTACACCAACCGAGTCGGGACAGGAGCC 960  
DB 1023 GTAGCCAGTCAATGGTGGCTCAACTGGGTTTACACCAACCGAGTCGGGACAGGAGCC 1082  
QY 961 CTGCGTGAGGTGGGCAACCTCTAGTGTGACCTTTGGCCGGGCTCCATGGGGACTGGTTA 1020  
DB 1083 CTGCGTGAGGTGGGCAACCTCTAGTGTGACCTTTGGCCGGGCTCCATGGGGACTGGTTA 1142  
QY 1021 CGTGCCTCTCGCCGCCCTCTCTGGAGATGGAGCCCAAGGTGTGACGCTTCTCAAGCAG 1080  
DB 1143 CGTGCCTCTCGCCGCCCTCTCTGGAGATGGAGCCCAAGGTGTGACGCTTCTCAAGCAG 1202  
QY 1081 CAGTGCAGGAGGTGAAGCACTGTACCTGGGCTGGAGCGGCAACATCTCTTAAGAAG 1140  
DB 1203 CAGTGCAGGAGGTGAAGCACTGTACCTGGGCTGGAGCGGCAACATCTCTTAAGAAG 1262  
QY 1141 TCCGGAAGGAGATGATGGCTCAAAATCTGAAGGTGACCCGGCTGAGGATACCTCT 1200  
DB 1263 TCCGGAAGGAGATGATGGCTCAAAATCTGAAGGTGACCCGGCTGAGGATACCTCT 1322  
QY 1201 TCTGCCCTTGCAAGAGACGCTTTAAGTTCGGAAGGCAATTTCAAGAAAGTCTCTCT 1260  
DB 1323 TCTGCCCTTGCAAGAGACGCTTTAAGTTCGGAAGGCAATTTCAAGAAAGTCTCTCT 1382  
QY 1261 ACCTCGTCAGGGAGGTACAGGAGGACCTCAGAACTCAGACCGGTGCTGATCTCCA 1320  
DB 1383 ACCTCGTCAGGGAGGTACAGGAGGACCTCAGAACTCAGACCGGTGCTGATCTCCA 1442

QY 1321 GGGCAGCTGTCCCTGCTGTATCCCTGCTCCCAAGGAAAGGCATCTCTAAAGATCTCGA 1380  
DB 1443 GGGCAGCTGTCCCTGCTGTATCCCTGCTCCCAAGGAAAGGCATCTCTAAAGATCTCGA 1502  
QY 1381 CAGCGTGAATCTGTTACTACTCTCTCCAGAGCCAGCGAGTCTGGGAACTCTTAGAC 1440  
DB 1503 CAGCGTGAATCTGTTACTACTCTCTCCAGAGCCAGCGAGTCTGGGAACTCTTAGAC 1562  
QY 1441 GCGAGTCAATGTTGTGAGTGGGACCCGCTGGAGCAGAGTCTCCACAGGCTTCAGGG 1500  
DB 1563 GCGAGTCAATGTTGTGAGTGGGACCCGCTGGAGCAGAGTCTCCACAGGCTTCAGGG 1622  
QY 1501 CTCTCTCTCCACCAAGGCAATTTCTCAAACTCAATGGCAAGTTCTCCCGCAGCCCTTA 1560  
DB 1623 CTCTCTCTCCACCAAGGCAATTTCTCAAACTCAATGGCAAGTTCTCCCGCAGCCCTTA 1682  
QY 1561 GAAGGCACTACCTCTAGCAGCTTTGGCTCTCTGGACCAACTGGCTCTCTCCATCTGCA 1620  
DB 1683 GAAGGCACTACCTCTAGCAGCTTTGGCTCTCTGGACCAACTGGCTCTCTCCATCTGCA 1742  
QY 1621 GCGCGCCAGCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTGTCTCCGAGTCC 1680  
DB 1743 GCGCGCCAGCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTGTCTCCGAGTCC 1802  
QY 1681 TTTGACCAATTTGCACTTGGCTGAACGCTTTCCGAAACCCCACTGAGGGGCTGTGTCT 1740  
DB 1803 TTTGACCAATTTGCACTTGGCTGAACGCTTTCCGAAACCCCACTGAGGGGCTGTGTCT 1862  
QY 1741 GTGACAACTGAGGGGCTTTGAGCAGCCCTCCCTCAGAAAGTCTGAAGCGATGTGGCAG 1800  
DB 1863 GTGACAACTGAGGGGCTTTGAGCAGCCCTCCCTCAGAAAGTCTGAAGCGATGTGGCAG 1922  
QY 1801 GAATCTCTGGGGATAGCTCTTTCTCTGACAGACTGCCAAGAGTGACTGCGAGCCTAC 1860  
DB 1923 GAATCTCTGGGGATAGCTCTTTCTCTGACAGACTGCCAAGAGTGACTGCGAGCCTAC 1982  
QY 1861 AGACAAGCCCTAGGAATCTCTCTCAAAAGCTCAGC 1893  
DB 1983 AGACAAGCCCTAGGAATCTCTCTCAAAAGCTCAGC 2015

## RESULT 3

US-09-799-451-215  
; Sequence 215, Application US/09799451  
; Patent No. 6783969  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Zhang, Jie  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Ma, Yuning  
; APPLICANT: Yamazaki, Victoria  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Ghosh, Reena  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 803  
; CURRENT APPLICATION NUMBER: US/09/799,451  
; CURRENT FILING DATE: 2001-03-05  
; NUMBER OF SEQ ID NOS: 948  
; SOFTWARE: pt\_FL\_genes Version 2.0  
; SEQ ID NO 215  
; LENGTH: 2501

TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (106)..(1989)  
US-09-799-451-215

Query Match 71.68; Score 1355.4; DB 4; Length 2501;  
Best Local Similarity 84.08; Pred. No. 0;  
Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

47 CTTCCGCGCTCGGCTCGGAGCGCCGCGCGCTGGCGGACGGGCTCATCAAGTCGCGCTA 106  
140 CCACTCCTCGGCGCAGAGCTAGCCGCGCGCTGGCGGAGGGCTCATCAAGTCGCGCA 199  
107 AACCTCTGATGAAGAAGCAGCGGCTGAAGCGGCAACATCAAAACAACAACCTCGCGGACC 166  
200 AGCCCTTAATGAAGAAGCAGCGGCTGAAGCGGCAACCAACCAAGCAACCTCGCGGACC 259  
167 GCTACGAGTTCCTGGAGCGCTGGCAAGGCGCACTACGGGAAGGTGAAGAGCGCAGAG 226  
260 GCTACGAGTTCCTGGAGCGCTGGCAAGGCGCACTACGGGAAGGTGAAGAGCGCGGG 319  
227 AGAGCTCGGCGCTCGTGGCCATCAAGTCCATCAGGAAGAAGCAAAATCAAAAGATGAGC 286  
320 AGAGCTCGGCGCTCGTGGCCATCAAGTCCATCAGGAAGAAGCAAAATCAAAAGATGAGC 379  
287 AGAGTCTGCTGCATACGAGGAGATGAGATCATGTCTTCACTCAACACCCCAACA 346  
380 AAGATCTGATGCATACGAGGAGATGAGATCATGTCTTCACTCAACACCCCAACA 439  
347 TCATTGCCATCCATGAAGTGTTCAGAAATGACAGCAAGATGATGATGATGATGATG 406  
440 TCATTGCCATCCATGAAGTGTTCAGAAATGACAGCAAGATGATGATGATGATGATG 499  
407 CCAGCCGAGGCGATCTGTATGATTAATCATCAGTACGCGGCGCAACGCTGAGTACGCGGACG 466  
500 CCAGCCGAGGCGATCTGTATGATTAATCATCAGTACGCGGCGCAACGCTGAGTACGCGGACG 559  
467 CCAGGCAATTTCTCGGAGATGCTGTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 526  
560 CTAGGCAATTTCTCGGAGATGCTGTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 619  
527 TTCAACGAGATCTCAAGCTGGAACATCTCTTCTAGATGCCAATGGAACATCAAGATTG 586  
620 TCACCGAGATCTCAAGCTGGAACATCTCTTCTAGATGCCAATGGAACATCAAGATTG 679  
587 CTGACTTTGGCTCTCCAACTGTACCAAGCAAGTTCCTCCAGACGTTCTGTGGGA 646  
680 CTGACTTTGGCTCTCCAACTGTACCAAGCAAGTTCCTCCAGACGTTCTGTGGGA 739  
647 GCGCTCTCTACGCTCGCTGAGATGATCAAGGGAAGCCCTATGTGGGCCAGAGTGG 706  
740 GCGGCTCTATGCTCGCAGAGATGTTCAATGGGAAGCCCTACACAGGCGCCAGAGTGG 799  
707 ACAGCTGCTCTCGGCGTCTCTGTATACCTCTGATGCGGACCATGCGCTTTGACG 766  
800 ACAGCTGCTCTCGGCGTCTCTGTATACCTCTGATGCGGACCATGCGCTTTGATG 859  
767 GGCAAGATCATAAACACATGCTGTAAGCAAAATCAGTAAGCGGGCTTACCGTGAGCGGCCCA 826  
860 GGCAATGACCAATAGATCTAGTGAACAGATCAGCAAGCGGGCTTACCGGAGCCACCTA 919  
827 AGCGCTCGATGCTGCGCTGATGCTGCGTGGCTGTTAATGGTGAACCCCAACCGTGGG 886  
920 AACCTCTGATGCTGCGCTGATGCTGCGTGGCTGTTAATGGTGAACCCCAACCGTGGG 979  
887 CCACACTGGAGATGATGAGTCAATGCTGGGTCAACTGGGGTTACACCAACCGGAGTGG 946  
980 CCACCTGGAGATGATGAGTCAATGCTGGGTCAACTGGGGTTACACCAACCGGAGTGG 1039  
947 GGGAAACAGGAGCCCTGCTGAGGCTGGGCAACCTAGTGTGATCTTTGGCGCGGCTTCCA 1006

## RESULT 4

US-09-984-890-1

; Sequence 1, Application US/09984890

; Patent No. 6492156

; GENERAL INFORMATION:

; APPLICANT: YAN, Chunhua et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USBS

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL001306

1040 GAGAGCAGGAGGCTCCGATGAGGGTGGGCAACCTGGCAGTGAATCTGTGCGCGCGCTCCA 1099  
1007 TGGCGAGCTGGTTACGTGCTCTCTGCGCGCGCGCTCTCTGAGATGAGACCAAGGTGCA 1066  
1100 TGGCTGACTGGCTCCGCGCTTCTCCGCGCGCGCTCTCTGAGATGAGGGCCAAAGGTGCA 1159  
1067 GCTTCTTCAAGCAGACGCTGCGCGGAGGTGGAGAGCTGTACTCTGGGCTGGAGCGGCAAC 1126  
1160 GCTTCTTCAAGCAGATGACCTGGTGGGGAAGACACCCCTGCGCTGGAGCGCGCAGC 1219  
1127 ATTCTCTTAAAGAGTCCGAAAGAGAGATGACATGCTCAAAATCTGCAAGGTGACCGG 1186  
1220 ATTCTCTTAAAGAGTCCGCAAGAGATGACATGCTGCGCGCGCTCTCTCAAGTGAACGG 1279  
1187 CTGAGATACCTTCTTCTGCGCGCTGGCAAGCAGACCTTAAGCTTCCGAAAGGATCTCA 1246  
1280 CTGATGACACTGCGCATGCGCTTGGCAAGAGCAACCTCAAGCTGCGCAAGGGATCTCA 1339  
1247 AGAAAGTCTCTTACTCTGCTCAGGGAGGTACAGGAGACCTCAGGAACTCAGACCGG 1306  
1340 AGAAGAGGTGTGACGCTCTGCAAGAGGGGTACAGGAGACCTCCTCGAGCTCAGCCCA 1399  
1307 TGCTGATCTCAGGGAGCGCTGCTCCCTGCTGTATCCCTGCTTCCCAAGAAAGGATCC 1366  
1400 TCCCTGCGAGCCCGAGCGGCTGCGC-----CCCTGCTCCCAAGAGGCAATTC 1450  
1367 TTAAGAGTCTCGACAGCGTGAATCTGTTACTTACTCTCTCCAGAGCCAGCGAGTCTG 1426  
1451 TCAAGAGGCGCGGAGCGAGTCTGGCTACTTCTCTCCGAGCCCGAGTGAATCTG 1510  
1427 GGGAACTCTTAGACGCGCAGTGTGTTGTGAGTGGGACCCCTGAGGAGCAGAAAGTCTC 1486  
1511 GGGAGCTCTTGGACGCGAGCGAGCTGTTGTGAGTGGGATCCCAAGGAGCAGAGCCTC 1570  
1487 CAGAGGCTCAGGGCTCTCTCCACCGCAAGGCGATTTCTAACTCAATGCGAAGTTCT 1546  
1571 CGAAGCTTCAAGGCTGCTCTCTCATCGCAAGGCGATCTCTCAAACTCAATGCGAAGTTCT 1630  
1547 CCGCAGCAGCTTAGAAGGCACTACCCCTAGACACTTTGGCTCTCTGAGCAACTGGCCT 1606  
1631 CCAGAGAGCTTGGAGCTCGCGCGCGCGCGCGCGCTCAGGGCTGTGAGGAGCAGCATCC 1690  
1607 CTTCCCATCTCGAGCGCGCGCGCGCGCGCTCAGGGGCTGTGAGTGAAGCAGCATCC 1666  
1691 CACTCGCGCGCTGCGCGCGCGCGCGCGCGCTCAGGGCTGTGAGGAGCAGCATCC 1750  
1667 TGTCTCCGAGTCTTGTGACCAATTTGACTTGGCTGNAAGTCTTCCGAAACCCCACTGA 1726  
1751 TGTCTCTGAGTCTTGTGACCAATTTGACTTGGCTGNAAGTCTTCCGAAACCCCACTGC 1810  
1727 GGGGCTGTGTGTCTGTGGCAACCTGAGGGGCTTGAAGCAGCTCTCTCAGAAAG----- 1780  
1811 GGGGCTGTGTGTCTGTGGCAACCTCAGGGGCTTGAAGGAGCCCTCAGAGGGCGCTG 1870  
1781 -----GTCTGAAGCAGATGTTGCGAGGAATCTTGGGGGATAGTGTCTTTCTGACAG 1834  
1871 GAAGTGCCTGAGCGCTGCGCGCGAGATCTTTGGGGGACAGCTGTCTTTCCCTGACAG 1930  
1835 ACTGCCAAGAGTGAATGAGCGCTACAGCAAGCCCTAGGAATCTGCTCAAGGCTCAGC 1893  
1931 ACTGCCAAGAGTGAATGAGCGAGCCTTACCGCAGGCACTGAGGGTCTGTCTCAAGGCTCACC 1989



```
; Sequence 1257, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1257
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1257

Query Match          10.0%; Score 189.2; DB 4; Length 484;
Best Local Similarity 64.3%; Pred. No. 6.2e-44;
Matches 284; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

Qy 233 CGGGGCTCTGGTGGCCATCAAGTCCATCAGGAAGAAGCAAAATCAAGATGAGCAGGATC 292
Db 442 CCGGCCAGGAGGTGGCTATCAAAACCATCAAGAAAGTCAAGATCGAGGCCGAGCGGATT 383
Qy 293 TGTGTCACATACGAGGAGGAGATTGAGATCATGTCTTCACTCAACACCCCCACATCATTTG 352
Db 382 TGGTGGCATCCGTCGGAGGTCAGATTATGAGCTCAGTGCATCATCCCAATCATATCC 323
Qy 353 CCATCCATGAAGTGTGTTGAGAAATAGCAGCAAGATTGTGATTTGTCTGAGAGTATGCCAGCC 412
Db 322 ACATCTACGAAGTATTGAGAAATCGTGAGAAATGGTGTCTAGTCTAGTCAATGAAATTTGCCGCTG 263
Qy 413 GAGGCGATCTGTATGATTATCATCAGTGCAGCGGCCACGGTGTGAGTGCAGGAGGCCAGGC 472
Db 262 GCGGCGAGCTCTACGACTATCTGTCTGAAGAGGAGGTTCTCACGAGGAGGAGCGGAGAC 203
Qy 473 ATTTCTTCCGACAGATCGTGTCTGCCCTGCCTACTCTGCCACCAAGACGGGATGTTCCACC 532
Db 202 GCATCTTCCGCGAGGTGGCCACCGCGTCTACTGTCTACAAGCAAGATCTGCCATC 143
Qy 533 GAGATCTCAAGCTGGAAACATCTTCTAGATCCCAATGGAAACATCAAGATTGCTGACT 592
Db 142 GCGATCTCAAGCTGGAGAACATCTCTGTGAGAGGGAATGCTTAAGATTGCTGATT 83
Qy 593 TTGGCCTCTCAACCTGTACCAAGGCAAGTTCTCTCAGAGCTTCTGTGGAGCCCTC 652
Db 82 TTGGGTTGTCAATGTGTTGATGACGAGCATGCTTGGGCACTTTTGGGGTTTCCCCAC 23
Qy 653 TCTACGCTCGCCTGAGATAGT 674
Db 22 TCTATGCTCGCCGAAATGTT 1

RESULT 8
US-09-930-181-1
; Sequence 1, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 V1
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(2112)
US-09-930-181-1

Query Match          9.8%; Score 186.4; DB 4; Length 2908;
Best Local Similarity 53.5%; Pred. No. 1e-42;
Matches 413; Conservative 0; Mismatches 356; Indels 3; Gaps 1;

Qy 166 CGGTACGAGTCTCTGGAGACGCTGGGCAAGGCAACCTTACGCGAAGGTGAAGAAGGCACGA 225
Db 157 CCCTACCGGTGGAGAAGACGCTGGGCAAGGCGAGACAGTCTGTGTGAAGCTTGGGGGTT 216
Qy 226 GAGAGCTCGGGGCGTCTGGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAG 285
Db 217 CACTGCGTCACTCCCAAGAGGTGGCCATCAAGATCGTCAACCGCTGAGAGCTCAGCGAG 276
Qy 286 CAGGATCTGTGCACATACGAGGAGATTGAGATCATGTCTTCACTCAACCCACCCAC 345
Db 277 TCGGTGTGATGAAGGTGGAGCGGAGATCCGATCTGAAGCTCATTTGAGCACCCCCAC 336
Qy 346 ATCATGCCATCCATGAAGTGTGTTGAGAAATAGCAGCAAGATTGTGATTGTGATGAGTAT 405
```



Db 337 GTCTAAAGTCGACGAGCTTTATGAAAAAATAATTTGTACCTGGTCTAGAACAC 396  
Qy 406 GCACGCCGAGCGATCTGTATGATTACATAGTCAGCGGCCACCGCTGAGTGCGGGAC 465  
Db 397 GTGTACGGTGGTGAAGCTCTTCGACTACTCTGTGAAGGGGAGGCTGACGCCCTAAGGAG 456  
Qy 466 GCCAGGCAATTTTCCGACAGATCGTGTCTGCCCTGCACACTGTCACCCAGAACCGGATC 525  
Db 457 GCTCGAAGTTCTTCCGCGAGATCATCTCTGCGCTGGACTTCTGCCACAGCCACTCCATA 516  
Qy 526 GTTCACCGAGATCTCAAGCTGGAACATCTCTTAGATGCAATGGAATCAAGATT 585  
Db 517 TGCCACAGGGATCTGAACCTGAAACCTCTCTGCTGACGAGAAAGAACATCCGCATC 576  
Qy 586 GCTGACTTTGGCTCTCTCAACCTGTACACAAAGGCAAGTTTCTTCCAGAGCTTCTGTGGG 645  
Db 577 GCAGACTTTGGCTGGCTGCTGCGAGGTGGCGACAGCTTGTGGAGACAGCTGTGGG 636  
Qy 646 AGCCTCTCTACGCTCGCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCACAGAGTG 705  
Db 637 TCCGCCCACTACGCTGCCCGAGGTGATCCGGGGGAGAAATATGACGGCGGGAAGCGC 696  
Qy 706 GACAGCTGGTCTCTGGGGGTTCTCTGTACATCTCTGTGATGCGCACCATGCCCTTTGAC 765  
Db 697 GACGTGTGGAGCTGCGGGCTCATCTCTGTCTGCTGTGGGGGCTCTGCCCTTCGAC 756  
Qy 766 GGGCAGGATCATAAACACACTGTTGAAGCAAAATCAGTAACGGGGCTTACCGTGAGCGCC 825  
Db 757 GATGACAACTTGGACAGCTGCTGGGAAGGTGAAGCGGGCGTGTTCACATGCCCGAC 816  
Qy 826 ---AAGCGTCCGATGCTGTGGCTGATCCGGTGGCTGTTAATGGTGAACCCCGCCGT 882  
Db 817 TTTATCCCGCGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGAGCGCGCAGC 876  
Qy 883 CGGCGCACACTGGAGGATAGCCAGTCATTTGGTGGGTCAACTGGGGTTACA 934  
Db 877 CGCCTCAGCTAGACACATTCAGAAAACACATATGTTATAGGGGGCAAGA 928

RESULT 9  
US-08-557-006C-39  
; Sequence 39, Application US/08557006C  
; Patent No. 6258547  
; GENERAL INFORMATION:  
; APPLICANT: Beri, Rajindar K.  
; APPLICANT: Carling, David  
; APPLICANT: Forster, Robert A.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
; FILE REFERENCE: NGAP/PHM37588/UST  
; CURRENT APPLICATION NUMBER: US/08/557,006C  
; CURRENT FILING DATE: 1996-03-06  
; PRIOR FILING DATE: 1994-05-20  
; PRIOR FILING DATE: 1993-05-21  
; PRIOR FILING DATE: 1993-08-31  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 39  
; LENGTH: 2652  
; TYPE: DNA  
; ORGANISM: Rat  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (1)..(1747)  
; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -  
; OTHER INFORMATION: fragment begins at nucleotide 24 and ends with  
; OTHER INFORMATION: nucleotide 1765  
US-08-557-006C-39

Query Match 9.6%; Score 182.4; DB 3; Length 2652;

Best Local Similarity 53.0%; Pred. No. 1.3e-41;  
Matches 437; Conservative 0; Mismatches 381; Indels 6; Gaps 2;  
Qy 120 GAACGAGCGGTGAAGCGGACCATCACAACACCAACCACTCGGCGCACCTACGAGTTCT 179  
Db 4 GAACATGGCTGGAAGCAGAACGACGACGGCGGTGTAAGATCGGACACTACCTGCTGGG 63  
Qy 180 GGAGACGCTGGGCAAGGGCACCTACGGGAAAGGTGAAGAGCGACGAGAGAGCTCG---GG 236  
Db 64 GGACACCTCGGGCGTGGCACCTTCGGCAAGTGAAGATTGGAGAAATCAATTGACAGG 123  
Qy 237 GCGTCTGGTGGCCATCAAGTCCATCAGAAAGCAAAATCAAAAGATGAGCAGGATCTGCT 296  
Db 124 CCATAAGTGGCATTAAAGATCTTAATAGACAGAGATTCCGAGTTTAGATGTTGTGG 183  
Qy 297 GCACATACGGAGGAGATTCAGATCATGTCTTCACTCAACCAACCCACATCATTTGCCAT 356  
Db 184 AAAAAATAAACGAGAAATCAAAATCTTTAAACTCTTTTCGTCATCTCATATATCAAACT 243  
Qy 357 CCATGAAGTGTTCAGAAATAGCAGCAAGATTGTGATTGTTCATGGAGTATGCCAGCCGAG 416  
Db 244 CTACCAAGTGTATCAGCACTCCACAGACTTTTTTATGTTAATGGAATATGTGCTGAGG 303  
Qy 417 CGATCTGTATGATTACATCAGTGAAGGCCCAACGGCTGAGTGAGCGGACCGCAGGCAATTT 476  
Db 304 TGAATTGTTCCGACTACATCTGTAAACACAGGGAGGTTGAAGAGGTGAAGCTGCCCGCT 363  
Qy 477 CTTCCGACAGATCGTGTCTGCCCTGACACTACTGCCACCAACCGGATCGTTTACCAGGA 536  
Db 364 CTTCCGACAGATTTCTGTCTGCCGTGGACTACTGTCTCAGGACATATGGTTTGTCCACAGGA 423  
Qy 537 TCTCAAGCTGGAACCAATCTCTTAGATGCAATGGAACATCAAGATTGCTGACTTTGG 596  
Db 424 CTTGAAGCCAGAGAACTGTGTCTGACGCCAGATGATGCTAAGATAGCTGACTTGG 483  
Qy 597 CTTCTCCAACTGTACCAAAAGCAAGTTCTTCCAGACGTTCTGTGGGAGCCCTCTCTA 656  
Db 484 ACTCTCTAATATGATGTCAGATGGTGAATTTCTACGAACTAGCTGTGATCGCCAAATTA 543  
Qy 657 CGCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCAGAGGTGACAGCTGGTC 716  
Db 544 TGACGACCGGAGGTCTATCTCAGGAAAGGCTGTATGCGGGTCTCTGAGTTGATATCTCGAG 603  
Qy 717 TCTGGGCTTCTCTGTACATCTCTGTGATCGGACCATCGCCCTTTGACGGGAGGATCA 776  
Db 604 CTGTGGTGTATCTGTATGCCCTTCTCTGTGGCACCTCCGTTCCAGCATGAGCAGT 663  
Qy 777 TAAACACATGGTGAAGCAAAATCAGTAACGGGGCTTACCGTGAAGCGGCCCGGTCGGA 836  
Db 664 GCCTACGCTCTTTAAGAAAGATCCGAGGGGTGTGTTCTACATCCCGAGTATCTCAACCG 723  
Qy 837 TGCTGTGGC---CTGATCGGTGGCTGTTAATGGTGAACCCACCCCGTGGGCCACACT 893  
Db 724 TTTCTATTGCCACTCTCTGATGCACATGCTGCGAGGTGGACCCCTTGAAGCGAGCAACTAT 783  
Qy 894 GGAGGATGTAGCCAGTCATTGGTGGGTCAACTGGGGTTACACCA 937  
Db 784 CAAGACATACGAGAGCATGAATGGTTTAAACAGGATTTGCCCA 827

RESULT 10  
US-08-557-006C-24  
; Sequence 24, Application US/08557006C  
; Patent No. 6258547  
; GENERAL INFORMATION:  
; APPLICANT: Beri, Rajindar K.  
; APPLICANT: Carling, David  
; APPLICANT: Forster, Robert A.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
; FILE REFERENCE: NGAP/PHM37588/UST  
; CURRENT APPLICATION NUMBER: US/08/557,006C  
; CURRENT FILING DATE: 1996-03-06  
; PRIOR FILING DATE: 1994-05-20  
; PRIOR FILING DATE: 1993-08-31  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 24  
; LENGTH: 2652  
; TYPE: DNA  
; ORGANISM: Rat  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (1)..(1747)  
; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -  
; OTHER INFORMATION: fragment begins at nucleotide 24 and ends with  
; OTHER INFORMATION: nucleotide 1765  
US-08-557-006C-24

;; PRIOR FILING DATE: 1994-05-20  
;; PRIOR APPLICATION NUMBER: GB 9310489.1  
;; PRIOR FILING DATE: 1993-05-21  
;; PRIOR APPLICATION NUMBER: GB 9318010.7  
;; PRIOR FILING DATE: 1993-08-31  
;; NUMBER OF SEQ ID NOS: 44  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 24  
;; LENGTH: 2761  
;; TYPE: DNA  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: cDNA of rat  
;; OTHER INFORMATION: liver AMP protein kinase  
US-08-557-006C-24

Query Match 9.6%; Score 182.4; DB 3; Length 2761;  
Best Local Similarity 53.0%; Pred. No. 1.3e-41;  
Matches 437; Conservative 0; Mismatches 381; Indels 6; Gaps 2;  
QY 120 GAAGCAGCGGTGAAGCGGCACCATCAAAACAACTCGGGCAGCGCTACGAGTTCTT 179  
DB 20 GAACATGCTGAGAAGCAGACGACGCGGTGTGAAGATCGACACTACGTGCTGG 79  
QY 180 GGAGACGCTGGCAGGCAAGGCACCTACGGAAGGTGAAGAGCCAGCAGAGAGCTGG 236  
DB 80 GGACACCCCTGGCGCTGGCACCTTCGGCAAGTGAAGATTGGAGAACATCAATTGACAG 139  
QY 237 GCGTCTGTGGCCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGCAGGATCTGCT 296  
DB 140 CCATAAAGTGCGAGTTAAGATCTTAATAGACAGAGATTCGCAAGTTAGATGTTGG 199  
QY 297 GCACATACGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCACATCATTCGCAT 356  
DB 200 AAAAAATAAACGAGAAATTCAAAATCTTAAACTCTTTCGTATCTCTATATTAACAAC 259  
QY 357 CCATGAGTGTGTGAGATAGCAGCATGATGATGATGATGATGATGATGATGATGATGAT 416  
DB 260 CTACCAAGTATCAGCATCTCCACAGACTTTTTATGTTAATGAATATGATGATGATGATG 319  
QY 417 CGATCTGTATGATTACATCAGTGAAGCGCCACGCTGAGTGAAGCGGACGCGCAGCATTT 476  
DB 320 TGAATTGTCAGTACATCTGTAAACACGGGAGGTTGAAGAGGTGAAGCTCCCGGCT 379  
QY 477 CTTCCGACAGATCGTGTCTGCGCTGACTACTGACCAAGACGCGGATCTGTCACCGAGA 536  
DB 380 CTTCCAGCAGATTTCTGTCTGCGGTGACTACTGTACAGGACATGTTGTCCACAGGGA 439  
QY 537 TCTCAAGCTGGAAACATCTTCTAGATGCCAATGGAACATCAAGATTGCTGACTTTGG 596  
DB 440 CTTGAAGCCAGAAACGTTGCTGCGACGCCAGATGATGCTAAGATAGCTGACTTCGG 499  
QY 597 CCTCTCCAACTGTACCAAGCAAGTTCTCCAGACGTTCTCCAGACGTTCTGTGGAGCCCTCT 656  
DB 500 ACTCTCTAATATGATGTAGATGTTGAATTTCTACGAACTAGCTGTGATCGCCAAATTA 559  
QY 657 CGCCTCGCTGATAGTCAAGCGGAGCCCTATGTGGCCGACAGAGTGGACAGCTGTGTC 716  
DB 560 TGCAGCAGCGGAGGTCTCTCAGGAAGGCTGTATCGGGTCTGAGGTTGATATCTGGAG 619  
QY 717 TCTGGGCTTCTCTGTATACCTCTGTGATGGCACCATTGCGCTTTGACGGGACGATCA 776  
DB 620 CTGTGGTGTATCTGTATGCTCTCTGTGCGACCCCTCCCGTTTCGAGGATGACACGT 679  
QY 777 TAAAACTGTGGAAGCAAAATCAGTAAACGGGCTTACCGTGAAGCCGCGCAAGCCGTCGA 836  
DB 680 GCCTACGCTCTTTAAGAGATCCGAGGGGTGTGTTCTACATCCCGGAGTATCTCAACCG 739  
QY 837 TGCCTGTGCG---CTGATCCGCTGGTGTATATGTTGAACCCACCCCTCGGGCCACACT 893  
DB 740 TTCTATTGCACTCTGCTGATGACATGCTGAGGTGGACCCCTTGAAGCGAGCAACTAT 799  
QY 894 GGAGGATGTAGCCAGTCAATTGGTGGTCAACTGGGGTTACACCA 937

DB 800 CAAAGACATACGAGACATGAATGTTTAAACAGGATTTGCCCA 843  
RESULT 11  
US-08-557-006C-38  
; Sequence 38, Application US/08557006C  
; Patent No. 6258547  
; GENERAL INFORMATION:  
; APPLICANT: Beri, Rajindar K.  
; APPLICANT: Carling, David  
; APPLICANT: Forder, Robert A.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
; FILE REFERENCE: NGAP/PHM37588/UST  
; CURRENT APPLICATION NUMBER: US/08/557,006C  
; CURRENT FILING DATE: 1996-03-06  
; PRIOR APPLICATION NUMBER: PCT/GB94/01093  
; PRIOR FILING DATE: 1994-05-20  
; PRIOR APPLICATION NUMBER: GB 9310489.1  
; PRIOR FILING DATE: 1993-05-21  
; PRIOR APPLICATION NUMBER: GB 9318010.7  
; PRIOR FILING DATE: 1993-08-31  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 38  
; LENGTH: 1742  
; TYPE: DNA  
; ORGANISM: Human AMP protein kinase  
US-08-557-006C-38

Query Match 9.6%; Score 181.2; DB 3; Length 1742;  
Best Local Similarity 53.1%; Pred. No. 2.3e-41;  
Matches 434; Conservative 0; Mismatches 378; Indels 6; Gaps 2;  
QY 126 GCGGTGAAGCGGCACCATCAAAACAACTCGGGCAGCGCTGTAAGATCGGACACTACGTCTGGGGACAC 185  
DB 3 GGTGAGAAAGCAGACGACGCGGCTGTGAAGATCGGACACTACGTCTGGGGACAC 62  
QY 186 GCTGGCAAGGGCACCCTACGGGAAGGTGAAGAGGCACGAGAGAGCT---CGGGCGTCT 242  
DB 63 CTTGGCGTTCGGCACCTTCGGCAAGTGAAGATTGGAGAACAATCAATTGACAGGCCATTA 122  
QY 243 GGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGCAGGATCTGCTGCACAT 302  
DB 123 AGTGGCAGTTAAGATCTTAAATAGACAGAGATTTCGAGTTTAGATGTTGTTGGAATAAT 182  
QY 303 ACGGAGGAGATTGAGATCATGTCTTCACTCAACACCCCACTCATTTGCCATCCATGA 362  
DB 183 AAAACGAGAAATCAAAATCTTAAACTCTTTCGTATCTCATATTTCAAACTTACCA 242  
QY 363 AGTGTGTTGAGATAGCAGCAAGATTGTGATTTGTCATGGAGTATGCGAGCGAGCGGATCT 422  
DB 243 AGTATCAGCACTCCAAACAGACTTTTATGGTAAATGGAATATGTCTGAGGTTGAAT 302  
QY 423 GTATGATTACATCAGTGAAGCGGCCACGCTGAGTGAAGCGGACGCGGACATTTCTTCG 482  
DB 303 GTTTCGACTACATCTGTAAACACGCGGAGGTGAAGAGGTGAAGAGTCTGCGGCTCTTCCA 362  
QY 483 ACAGATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542  
DB 363 GCAGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422  
QY 543 GCTGGAAAAACATCTTCTAGATGCGCAATGGAACATCAAGATTGCTGACTTTGGCTCTC 602  
DB 423 GCCAGAGAACGTGTTGCTGGACGCCAGATGAAATGCTTAAGATAGCTGACTTCGAGCTCTC 482  
QY 603 CAACCTGTACCAAAAGCAAGTTCTCCAGAGTTCTGTTGGAGGCCCTCTCTACGGCTC 662  
DB 483 TAATATGATGTCAGATGTTGAAATTTCTACGAACTAGCTGTTGGATCGCAAAATTTATGAGC 542  
QY 663 GCCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCACAGAGGTGGACAGTGGTCTCTGGG 722  
DB 543 ACCGAGGTCTATCTCAGGAAGGCTGTATGCGGGTCTCTGAGGTTGATATCTGAGCTGTGG 602



```
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 112
; LENGTH: 2899
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(2899)
US-09-774-528-112

Query Match          9.3%; Score 176.4; DB 4; Length 2899;
Best Local Similarity 52.4%; Pred. No. 7e-40;
Matches 412; Conservative 0; Mismatches 371; Indels 3; Gaps 1;

QY 113 TGATGAAGAGCGCGGTGAAGCGGCACCATCACAAACAAACCTGCGGCACCGCTACG 172
Db 120 TGGTCATGGCGGTGGCCCGAGGCACCTTGACGCGCGCGGTCCGGGTGGGTCTCTACG 179

QY 173 AGTTCCTGGAGACGCTGGGCGAGGCGACCTACGGGAAGGTGAAGAGGACGAGAGCT 232
Db 180 ACATCGAGGCGACGCTGGGCGAGGCGAACTTCGCTGTGTGAAGTGGGCGGCGACCGGA 239

QY 233 CGGGCGCTCTGTGGCCATCAAGTCCCATCAGGAAGACAAAATCAAAAGATGAGCAGATC 292
Db 240 TCACCAAGACGGAGGTGGCAATAAATAATTCGATAAGTCTCAGCTGATGCGAGTGAAC 299

QY 293 TGTGTCACATACGGAGGAGATTGAGATCATGTCTTCACTCAACACCCCGACATCATTTG 352
Db 300 TTGAGAAAATCTACCGAGAGTACAAATAATGAAAATGTTAGACACCCCTCACATAATCA 359

QY 353 CCATCCATGAAGTGTGTGAGNATAGCAGCAAGATTGTGATGTCATGAGTATGCCAGCC 412
Db 360 AACTTTATCAGGTAAATGGAGACCAAAAGATGTTGTACCTTGTGACAAATATGCCAAA 419

QY 413 GAGCGGATCTCTATGATTACATCAGTGAGCGGCCACCGCTGAGTGAGCGGACGCCAGGC 472
Db 420 ATGAGAGAAATTTTGAATCTCTGCTAATCATGCGCGGTAAATGATCTGAAGCCAGGC 479

QY 473 ATTTCCTCGACAGATCGTGTCTGCCCTGCACTACTGCGCACAGAAACATCAAGATTGCTGACT 532
Db 480 GAAAATTTGCGCAAAATCTGTCTGCTGTGATTATTTGTATGTCGGAAGATTGTGCACC 539

QY 533 GAGATCTCAAGCTGGAACATCTCTAGATGCCAATGGAAACATCAAGATTGCTGACT 592
Db 540 GTGACCTCAAGCTGGAACATCTCTGCTGGATTAACAACATGAATATCAAAATACAGATT 599

QY 593 TTGGCCTCTCAACCTGTACCAAGGCAAGTTCTCTCCAGACGTTCTGTGGGAGCCCTC 652
Db 600 TCGGTTTGGGAAATTTCTTTAAAGTGGTGAACCTGTGGCAACATGGTGTGGCAGCCGCC 659

QY 653 TCTAGCCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCGAGAGTGACAGCT 712
Db 660 CTTATGACGCCCGCAGAACTCTTTGAAGGGCGAGCAGTATGAAGGACCAACAGCTGGACATCT 719

QY 713 GGTCTCTGGCGGTTCTCTGTACATCTCTGTGTCATGGCACCATGCCCTTTTACCGGGCAGG 772
Db 720 GGAGTATGGGAGTTGTTCTTTATGTCCTGTGCTGTGAGAGCTCTGCCCTTTGATGACCGA 779

QY 773 ATCATAAACACATGCTGAAGCAAAATCAGTAACGGGGCTTACCGTGAGCGCGCCCAAGCCGT 832
Db 780 CTCCTTCCAATTTTGAAGGCGAGGGGTTCTGGAAGGAAGATTCGGATTCGGTATTTCTATGT 839

QY 833 CGATGCTCTGTG---GCCTGATCCGGTGGCTGTTAATGTTGAACCCCAACCGCTCGGGCCA 889
Db 840 CAGAAAGATTGCGAGCACCTTATCCGAAGAGATGTTGGTCTTAGACCCATCCAAACGGCTAA 899

QY 890 CACTGG 895
Db 900 CCATAG 905

RESULT 14
US-09-930-181-3
```

```
; Sequence 3, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: OxiGene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 V1
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3364
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (482)..(2239)
US-09-930-181-3

Query Match          9.2%; Score 174; DB 4; Length 3364;
Best Local Similarity 53.4%; Pred. No. 3.6e-39;
Matches 414; Conservative 0; Mismatches 355; Indels 7; Gaps 2;

QY 166 CGTACGAGTTTCTTGAGACGCTGGGCAAGGCGACCTACGGGAAGGTGAAGAGGACGCA 225
Db 280 CCTACCGGCTGGAGAAGACGCTGGGCAAGGCGGACAGAGGCTCTGGTGAAGCTGGGGGTT 339

QY 226 GAGAGCTCGGGGCTGTGTGGCCATCAAGTCCATCAGGAAAGACAAAATCAAGATGAG 285
Db 340 CACTGGCTCACCTGCCAGAAAGGTGGCCATCAAGATCGTCAACCGTGAGAGCTCAGCGAG 399

QY 286 CAGGATCTCTGCACATACGGAGGGAGATTGAGATCATGTCTTCACTCAACACCCGCCAC 345
Db 400 TCGTCTCATGAGGTGGAGCGGGAGATCGCATCTCTGAAGCTATTGAGCACCCGCCAC 459

QY 346 ATCATTTGCCATCCATGAAGTGTGAGAAATAGCAGCAAGATTGTGATT-----GTCAATGA 401
Db 460 GTCTTAAAGCTGCACGACGTTTATGAAAACAAAAAATATTTGTAGGTACTGTGTGTAGTA 519

QY 402 GTATGCCCGCGAGGCGATCTGTATGATTACATCAGTGAGCGGCCACGCTGAGTGAGCG 461
Db 520 ACAGTGTTCAGGTGGTGAGCTTTCGACATACCTGTTGTAAGAGGGGAGGCTGACGCTAA 579

QY 462 GGACGCGCAGGCATTTCTTCGACAGATCGTGTCTGCCCTGCACTACTGCGCACAGAAACGG 531
Db 580 GGAGGCTCGGAAGTTCTTCGGGCGAGATCATCTCGGCTGGACTTCTGCGCACAGCCACTC 639

QY 522 GATCGTTTACCGAGATCTCAAGCTGGAAACATCTCTTAGATGCCAATGGAAACATCAA 581
Db 640 CATATGCCACAGGGATCTGAAACCTGAAACCTCTGCTGGACGAGAAAGAACATATCCG 699

QY 582 GATTGCTGACTTTTGGCTCTCCAACTGTACCAAGGCAAGTCTCTCCAGACGTTCTG 641
Db 700 CATCGCAGACTTTTGGCATGGCTCCCTTCAGAGTTGGCGACAGCTGTTGGAGACAGCTG 759

QY 642 TGGGAGCCCTCTCTACGCTTCCGCTGAGATGATGTAACGGGAAGCCCTATGTGGGCCGAGA 701
Db 760 TGGTCCCGCCACTACGCTTCCCGCTGCCCGAGGTGATCGGGGGGAGAAAGTATGACGCGCGAA 819

QY 702 GGTGGAAGCTGTCTCTGGCGGTTCTCTGTGATACCTCTGTCATCGTCGACCAATGCCCTT 761
Db 820 GGGGACGCTGTGGAGCTGGCGGCTCATCTGTTGCGCTTCTGTTGGGGGGCTCTGCGCTT 879

QY 762 TGACGGGCGAGGATCATAAAACACTGGTGAAGCAAAATCAGTAACGGGGCTTTACCGTGAGCC 821
Db 880 CGACGATGA CAATTGCGACAGCTGTCTGGAGAGGTGAAGCGGGCGCTGTTCCACATGCC 939

QY 822 GCCC---AAGCGGTCCGATGCTGTGGCCCTGATCCGGTGGCTGTTAATGTTGAGACCCAC 878
Db 940 GCATTTTATCCGCCCCGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGAGCGCGC 999

QY 879 CCGTCGGGCCACACTGGAGGATGTAGCCAGTCAATTGGTGGTCAACTGGGGTTACA 934
```

```
Db 1000 AGCGCGCTCACGCTAGAGCACATTTCAGAAACACATATGGTATATAGGGGCAAGA 1055
RESULT 15
US-10-003-690-3
; Sequence 3, Application US/10003690
; Patent No. 6787345
; GENERAL INFORMATION:
; APPLICANT: RORY A.J. CURTIS
; TITLE OF INVENTION: 55053, A No. 6787345el Human Eukaryotic Kinase
; FILE OF INVENTION: and Uses Therefor
; FILE REFERENCE: MNI-206
; CURRENT APPLICATION NUMBER: US/10/003.690
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/248,893
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2334
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2334)
US-10-003-690-3

Query Match 9.1%; Score 173; DB 4; Length 2334;
Best Local Similarity 52.7%; Pred. No. 5.8e-39;
Matches 399; Conservative 0; Mismatches 355; Indels 3; Gaps 1;

Qy 165 CCGCTACGAGTTCCTGGAGACGCTGGGCAAGGGCACCTACGGGAAGGTGAAGAGGCACG 224
Db 96 CCCCTATCGGCTGGAGAGACGCTGGGCAAGGACACAGAGGCTGGTTAACTCGGGGT 155

Qy 225 AGAGAGTCGGGGCGTCTGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGA 284
Db 156 CCACTGCATCAGGGTCAGAAGGTCGCCATCAAGATCGTGAACGGGAGAAGCTGTCGA 215

Qy 285 GCAGGATCTGTCACATACGGAGGAGATTGAGATCATGTCTTCACTCAACACCCCA 344
Db 216 GTGGGTCTGATGAAGGTGGAGCGGGAGATCGCCATCTTGAAGCTCATCGAACACCCCA 275

Qy 345 CATCAITGCCATCCATCAAGTGTGTGAGATAGCAGCAAGATTGTGATGTGATGAGTA 404
Db 276 TGTCTCAAGCTCACAGCGCTACGAGAACAGAAATATTTGTACCTGGTTCTGGAGCA 335

Qy 405 TGCCAGCCGAGGGGATCTGTATGATTATCATCAGTGAGCGGCCACGGCTGAGTACGGGA 464
Db 336 CGTCTCGGGGGTGAGCTATTGAGTACCTGGTAAAGAGGGGAGACTGACGCCCAAGGA 395

Qy 465 GCGCAGGATTTCTTCGACAGATCGTGTGCGCTGCACATCTGCCACAGAACGGGAT 524
Db 396 GGGCCGAAGTCTTCGCGCAGATTTGTCTGCGCTGGACTTCTGCCACAGCTACTCCAT 455

Qy 525 CGTTACCGAGATCTCAAGCTGGAACATCTTCTAGATGCCAATGGAACATCAAGAT 584
Db 456 CTGCCACAGAGACCTAAAGCCGAGAACCTGCTTTTGGATGAGAAACAAACATCCGCAT 515

Qy 585 TGCTGATTTTGGCTCTCCAACTGTACCAAGGCAAGTTCTCCAGACGTTTCTGTGG 644
Db 516 TGCAGACTTCGGCATGGGCTCCCTGCGAGTGGGGGACAGCTCTCTGGAGACGAGTGGG 575

Qy 645 GAGCCCTCTCTACGCTCGGCTGAGATGATGACGGAAGCCCTATGTGGGCCAGAGGT 704
Db 576 GTCCCCCAATTTATCGGTGTCCAGAGGTGATTAAAGGGGAAAAATATGATGCGCGCGGC 635

Qy 705 GGACAGCTGGTCTCTGGGGCTTCTCCTGTACATCTGCTGATGACCATGCCCTTTGA 764
Db 636 AGACATGTGAGAGTGTGAGAGTCATCTCTTTCGCCCTCTCTGTGGGGGCTCTGCCCTTTGA 695

Qy 765 CGGCGAGGATCATAAACACTGGTGAAGCAAAATCAGTAAACGGGGCTTTACCGTGAAGCGGC - 823
```

```
Db 696 TGACGACAACTCCGCGAGCTGCTGGAGAAAGGTGAACCGGGCGCTTTCACATGCCCCA 755
Qy 824 --CCAAGCCGTCCGATGCGCTGTGCGCTGATCCCGTGGCTGTTAATGTTGAACCCACCG 881
Db 756 CTTCAATTCCTCCAGATTGCCAGAGCTTCTTGAGGGGAATGATCGAAGTGGAGCCCGAAA 815
Qy 882 TCGGGCCACACTGGAGGATGTACCCAGTCAITGGTGG 918
Db 816 AAGGCTCAGTCTGGAGCAAAATTCAGAAACATCCTTGG 852
```

Search completed: January 25, 2005, 01:59:45  
Job time : 156.572 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 05:56:58 ; Search time 875.636 Seconds  
(without alignments)  
11348.500 Million cell updates/sec

Title: US-09-980-464-4\_COPY\_123\_2015

Perfect score: 1893

Sequence: 1 atggagtcggtgccttact.....gaatcgtcctaagctcagc 1893

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1893	100.0	2902	4 AAC90433	Aac90433 Murine Ly
2	1683.4	88.9	2929	6 AAD14328	Aad14328 Rat SNF1
3	1360.2	71.9	1884	12 ADL14162	Adl14162 Novel hum
4	1360.2	71.9	3353	6 ABK14000	Abk14000 cDNA enco
5	1360.2	71.9	3353	12 ADL14160	Adl14160 Novel hum
6	1357.6	71.7	3360	6 AAD26459	Aad26459 Human kin
7	1357	71.7	3395	4 AAK94280	Aak94280 Human ful
8	1357	71.7	3395	12 ADL30885	Adl30885 Full leng
9	1357	71.7	3463	12 ADJ96554	Adj96554 Human cal
10	1355.4	71.6	2043	6 ABZ11333	Abz11333 Human pol
11	1355.4	71.6	2501	12 ADM43851	Adm43851 Novel hum
12	1355.4	71.6	3443	5 ABX71420	Abx71420 Human cel
13	1355.4	71.6	3443	10 ADF76964	Adf76964 Novel hum
14	1355.4	71.6	3443	10 ADF81952	Adf81952 Leukemia
15	1355.4	71.6	3443	12 ADO20171	Ado20171 Human PRO
16	1338.8	70.7	2291	4 AAF44659	Aaf44659 Novel pro
17	1338.8	70.7	2291	12 ADI29357	Adi29357 Human MAR
18	1314.2	69.4	3200	4 AAF75338	Aaf75338 Human TGF
19	1230	65.0	1833	10 ABZ77163	Abz77163 Human pro
20	840.2	44.4	2616	6 ABQ72599	Abq72599 Human MDD
21	840.2	44.4	2619	6 ABQ72698	Abq72698 Human MDD

22	506.4	26.8	2884	4 AAD14328	Aad14328 Human pro
23	506.4	26.8	6828	4 AAD14327	Aad14327 Human pro
24	506.4	26.8	6828	10 ADE38420	Ade38420 Human pro
25	506.4	26.8	6828	11 ADN95767	Adn95767 Human BEC
26	506.4	26.8	6828	12 ADJ74808	Adj74808 Marker ge
27	506.4	26.8	6828	12 ADL25352	Adl25352 Human ARK
28	506.4	26.8	6828	12 ADQ19733	Adq19733 Human sof
29	506.4	26.8	6828	12 ADP43253	Adp43253 Human pro
30	506.4	26.8	6854	12 ADO23883	Ado23883 Human sof
31	356.8	18.8	3594	5 AAO33994	Aao33994 Human pro
32	348.8	18.4	587	4 AAK93296	Aak93296 Human CDN
33	348.8	18.4	587	4 AAK91887	Aak91887 Human CDN
34	348.8	18.4	587	12 ADL29723	Adl29723 5' end of
35	348.8	18.4	587	12 ADL28314	Adl28314 5' end of
36	296.8	15.7	1454	4 ABL21269	Ab121269 Drosophil
37	267	14.1	1723	4 AAH13802	Aah13802 Human CDN
38	236.6	12.5	1594	4 AAF44655	Aaf44655 Novel pro
39	236.6	12.5	1594	8 AAL60326	Aal60326 Human 207
40	236.6	12.5	1594	12 ADI29353	Adi29353 Mouse MAR
41	236.6	12.5	3170	6 ABA05737	Abao5737 Murine ne
42	236.6	12.5	3250	6 ABA05738	Abao5738 Murine ne
43	233.6	12.3	1549	6 ABS64386	Abse64386 Human ser
44	233.6	12.3	2085	8 AAD51409	Aad51409 Human mic
45	233.6	12.3	2085	10 ADG91747	Adg91747 Human mic

ALIGNMENTS

RESULT 1  
AAC90433  
ID AAC90433 standard; cDNA; 2902 BP.

AC AAC90433;

XX 19-MAR-2001 (first entry)

XX Murine Lymph node Stromal cell kinase 1 coding sequence.

KW Murine; Lymph node Stromal cell kinase; MLK-1; autoimmune disorder;  
KW wound healing; peritoneal disease; inflammatory disease; tumour;  
KW infection; allergy; ss.

OS Mus musculus.

PN WO2000073468-A1.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US014696.

XX 28-MAY-1999; 99US-0136781P.

XX (IMMV ) IMMUNEX CORP.

XX Bird TA, Virca GD, Martin U, Anderson DM;

XX WPI; 2001-061546/07.

XX P-PSDB; AAB50056.

XX Novel murine and human kinase nucleic acids useful for treating  
PT inflammations, infections, tumors, allergies, autoimmune diseases, and  
PT for stimulating or suppressing immune responses.

XX Claim 1; Page 86-87; 106pp; English.

XX The present sequence is the coding sequence for Murine Lymph node Stromal  
CC cell kinase 1 (MLK-1). The protein encoded by the present sequence is  
CC useful for treating a variety of disorders listed in the disclosure of  
CC the specification, including autoimmune disorders, allergic reactions,  
CC myeloid or lymphoid cell deficiencies, wound healing and tissue repair,  
CC and replacement, burns, incisions and ulcers, periodontal disease,  
CC inflammatory diseases, tumours and bacterial, viral or fungal infection





Rat; SNF1/AMPK-Related Kinase; SNARK; enzyme; stress response; diabetes; glucose deprivation; lipid metabolism; therapy; lipoprotein disorder; hyperglycaemic; drug screening; hypoglycaemia; ss.

Rattus sp.

Key Location/Qualifiers  
CDS 83..1975  
/\*tag= a  
/product= "Rat SNARK protein"

WO200212456-A2.

14-FEB-2002.

02-AUG-2001; 2001WO-CA001109.

03-AUG-2000; 2000US-022650P.

12-MAR-2001; 2001US-0274613P.

28-MAR-2001; 2001CA-02340783.

(ONEO-) 1149336 ONTARIO INC.

Drucker DJ, Rosen CF, Lefebvre DL;

WPI; 2002-241747/29.

P-PSDB; AAE19885.

AMPK (AMP-activated protein kinase)-related kinase, designated SNARK polypeptides and polynucleotides, useful for treating or preventing diabetes, or other disorders of lipoprotein production leading to increased levels of cholesterol.

Example; Fig 2; 94pp; English.

The invention relates to an AMPK (AMP-activated protein kinase)-related kinase, designated SNARK polypeptides and polynucleotides. SNARK (SNIF/AMP-activated protein kinase) is involved in stress response to glucose deprivation. The polynucleotides are useful for expressing SNARK protein in isolated form or as a protein conjugate. Activation of SNARK stimulates liver CPT-1 thus enhances lipid metabolism in liver cells and in other cell types such as heart and skeletal muscles, as well as increases GLUT-4 and glycogen in muscle. Activation of SNARK is predicted to have insulin-like effects that would enhance the disposal of glucose into muscle and reduce plasma glucose for the treatment of diabetes and some type of disorders of lipoprotein production leading to increased levels of cholesterol or triglycerides. SNARK or its variants may be administered to a subject to treat or prevent a disease associated with decreased expression of SNARK, such as diabetes. SNARK antibodies are used to modulate SNARK activity either in vivo for therapeutic purposes, or in vitro, for drug screening and related investigational purposes. SNARK antagonists may be administered to increase fuel production, decrease glucose uptake and increase levels of blood glucose in a patient suffering from hypoglycaemia. The present sequence is rat SNARK cDNA

Sequence 2929 BP; 650 A; 788 C; 791 G; 700 T; 0 U; 0 Other;

Query Match 88.9%; Score 1683.4; DB 6; Length 2929;  
Best Local Similarity 93.4%; Pred. No. 0;  
Matches 1769; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

1 ATGAGTTCGGTGGCTTACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCCGCGCTGGCC 60  
83 ATGAGTTCGGTGGCTTACACCGCGGGGNAACCTGCTCCCTCGGCTCCGCGCTGGCC 142  
61 TCGAGAGCGCCCGCGCTGGCGGACGGGTCTCAATGCTGCTAAACCTCTGATGAAG 120  
143 ACGGAGAGCGCCCGCGCTGGCGGACGGGTCTCAATGCTGCTGCTGCTGATGAAG 202  
121 AAGCAGCGGTGAAGCGGCACCATCAAAACAACTGGGCAACCGGTAGGAGTTCCTG 180  
203 AAGCAGCGGTGAAGCGGCACCATCAAAACAACTGGGCAACCGGTAGGAGTTCCTG 262

181 GAGACGCTGGCGAAGGCACCTACCGGAAGGTGAAGAGGCAAGAGAGCTCGGGCGT 240  
263 GAGACCTTGGCAAGGCACTACGGGAAGGTGAAGAAAGCAAGAGAGCTCGGAGCG 322  
241 CTGCTGGCCATCAAGTCTCATCAGAAAGACAATAATCAAGATGAGCAGGATCTCTGCAC 300  
323 CTGCTGGCCATCAAGTCTATCAGGAAGGACAAAATCAAGATGAGCAGGATCTCTGCAC 382  
301 ATACGAGGAGAGTGAATCATCTTCACTCAACACCCACCATCATTTGCCATCAT 360  
383 ATAGGAGGAGATCGAGATCATGTCTTCACTCAACACCCACCATCATTTGCCATCAT 442  
361 GAAAGTGTTCAGAAATAGCAGCAAGATGTGATGTCTATGAGTATGCGAGCGAGCGAT 420  
443 GAAGTGTTCAGAAACAGCAGCAAGATGTGATGTCTATGAGTATGCGAGCGAGCGAT 502  
421 CTGTATGATTAATCATCAGTGAAGCGCCACGGCTGAGTGAAGCGGACGCGAGCATTTCTTC 480  
503 CTGTACGATTTACATCAGTGAAGCGCCACGGCTGAATGAGCGGACGCGAGCATTTCTTC 562  
481 CGACAGATCGTGTCTGCGCCCTGCACTACTGCGCACCAAGACGGGATCTTCAACGAGATCTC 540  
563 CGACAGATCGTGTCTGCGCCCTGCACTACTGCGCACCAAGACGGGATTTTCAACGAGATCTC 622  
541 AAGCTGGAACACATCTCTTAGATGCGCAATGGAACATCAAGATTTGCTGATTTGGGCTC 600  
623 AAGCTGGAACACATCTCTTAGATGCGCAATGGAACATCAAGATTTGCTGATTTGGGCTC 682  
601 TCCAACTGTACCAAAAGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTCTCTTAGGCC 660  
683 TCCAACTGTATCACAAAGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTCTCTTAGGCC 742  
661 TCGCTGAGATAGTCAACGGGAGCCCTATGTGGGCCCGAGGTTGGACAGCTGCTCTCTG 720  
743 TCACCTGAGATCGTCAACGGGAGCCCTATGTGGGCCCGAGGTTGGACAGCTGCTCTCTG 802  
721 GGGCTTCTCTGTACATCTCTGTGATGCGACCATGATGCGCTTTGACGGGAGGATCATAAA 780  
803 GGGCTTCTCTGTACATCTCTGTGATGCGACCATGATGCGCTTTGACGGGAGGATCATAAA 862  
781 ACATGCTGAAGCAATCAATACGGGCTTACCGTGAAGCGGCCCAAGCCGTCGATGCC 840  
863 ACCCTGTGAAACAAATCAGTAGCGGGCTTACCGAGAGCCGTGCAAAACCGCTCTGATGCC 922  
841 TGTGGCTGATCCGGTGGCTGTTAAATGGTGAACCCACCCGTCGGGCCACACTGGAGAT 900  
923 TGTGGCTGATCCGGTGGCTGTTAAATGGTGAATCCCATCCGTCGGGCCACTCTGGAGGAT 982  
901 GTAGCCAGTCAATTTGGTGGGTCAACTGGGGTTACACACCGGAGTTCGGGGAAACAGGAAGCC 960  
983 GTAGCCAGTCAATTTGGTGGGTCAACTGGGGTTACAGCACCCGAATTTGGGAAACAGGAAGCT 1042  
961 CTGCTGAGGCTGGGACCCCTAGTGTGACTTTTGGCGGCGCTCTCATGCGGAGTGTGTTA 1020  
1043 CTGCGAGGCTGGGACCCCTAGCGGTGACTCTGGCGGCGCTCTATGGCGGAGTGTGTTA 1102  
1021 CGTGGCTCTCGCGCCCTCTCTGGGAGATGGAGCCAAAGGTGTGACAGTCTTCTCAAGCAG 1080  
1103 CGTGGCTCTCTCGCGCCCTCTCTGGGAGATGGAGCCAAAGGTGTGAGTCTTCTCAAGCAG 1162  
1081 CAGTGGCGGAGGTGGAAGCACTGTACCTGGGCTGGAGCGGCAACAACTCTCTTAAGAG 1140  
1163 CATGTGGCGGAGGTGGAAGCACTGGGCTGGAGCGGCGCTGGAGCGGCAACAACTCTCTTAAGAG 1222  
1141 TCCGAAAGAGAGATGACATGGCTCAAAATCTGCAAGGTGACCGGCTGAGGATACCTCT 1200  
1223 TCCGAAAGAGAGATGACATGGCTCAAGTCTGAGAAATGACCCAGTGAAGATACCTCT 1282  
1201 TCTGCGCCCTGGCAAGAGACCTTTAGCTTCCGAAAGGCAATTTCTCAGAAAGAGTCTCT 1260  
1283 TCTGCGCCCTGGCAAGAGACCTTTAGCTTCCGAAAGGCAATTTCTCAGAAAGAGTCTCT 1342  
1261 ACCTCGTTCAGGAGGTACAGGAGGACCTCTCAGAACTCAGACCGGTCGCTGATCTCCA 1320

Db 1343 CCCTCATCGGGGAGGTACAGAGGGCCCTCAGGAATCTCAGACAGTGTCAATACCCCA 1402  
Qy 1321 GGGCAGCCTGTCTCTGTATCCCTGTCTCCCAAGGAAAGGCATCTTAAAGAGTCTCGA 1380  
Db 1403 GGGCAGCCTGTCTCTGTATCCCTGTCTCCCAAGGAAAGGCATCTTAAAGAGTCTCGG 1462  
Qy 1381 CAGCGTGAATCTGTTTACTTCTCTCCAGAGCCCAAGCGAGTGTGGGGAAGTCTTTAGAC 1440  
Db 1463 CAGCGTGAATCTGTTTACTTCTCTCCAGAGCCCAAGCGAGTGTGGGGAAGTCTTTAGAC 1522  
Qy 1441 GCCAGTGAATCTGTTTACTTCTCTCCAGAGCCCAAGCGAGTGTGGGGAAGTCTTTAGAC 1500  
Db 1523 GCAGGTGAATCTGTTTACTTCTCTCCAGAGCCCAAGCGAGTGTGGGGAAGTCTTTAGAC 1582  
Qy 1501 CTCCTCTCTCCAGAGCCCAAGGCATCTCAAACTCAATGCGAAGTCTCTCCGACAGCCTTA 1560  
Db 1583 ---CGCCTCCATCGCAGGGCATCTCAAACTCAATGCGAAGTCTCTCCGACAGCCTTA 1639  
Qy 1561 GAAGGCACTACCTCTAGACCTTTGGTCTCTCTGAGCAAACTGGGCTCTCTCCATCTCTGCA 1620  
Db 1640 GAAGGCACTACCTCTAGACCTTTGGTCTCTCTGAGCAAACTGGGCTCTCTCCATCTCTGCA 1699  
Qy 1621 GCCCGCCAGCCGCTCTGAGGAGTGTGAGTGAGGACAGCATCTCTCCGAGTCT 1680  
Db 1700 GCCCGCCAGCCGCTCTGAGGAGTGTGAGTGAGGACAGCATCTCTCCGAGTCT 1759  
Qy 1681 TTTGACCAATGGACTTGGCTTGAACCTCTCCGAAACCCCACTGAGGGCTGTGTCT 1740  
Db 1760 TTTGACCAATGGACTTGGCTTGAACCTCTCCGAAACCCCACTGAGGGCTGTGTCT 1819  
Qy 1741 GTGACAACTGAGGGCTTGGAGGCTCTCTGAGGAGTCTGAGGAGTGTGTCTGAG 1800  
Db 1820 GTGACAACTGAGGGCTTGGAGGCTCTCTGAGGAGTCTGAGGAGTGTGTCTGAG 1879  
Qy 1801 GAATCTTTGGGGATAGTCTTTCTCTGACAGACTCCCAAGAGTGTGAGGAGTGTGTCT 1860  
Db 1880 GAATCTTTGGGGATAGTCTTTCTCTGACAGACTCCCAAGAGTGTGAGGAGTGTGTCT 1939  
Qy 1861 AGACAAGCCCTAGGAATCTGTCAAGCTCAGC 1893  
Db 1940 AGACAAGCCCTAGGAATCTGTCTGAGGCTCAGC 1972

RESULT 3  
ID ADL14162 standard; cDNA; 1884 BP.  
XX AC ADL14162;  
XX DT 17-JUN-2004 (first entry)  
XX DE Novel human gene 3700 coding region.  
KW cytosolic; cardiac; hypotensive; antiangiogenic; osteopathic;  
KW antiarthritic; antirheumatic; neuroprotective; antiinflammatory;  
KW aniporiatic; antidiabetic; cardiovascular; virucide; analgesic; CNS;  
KW angogenesis inhibitor; angogenesis stimulator; cerebroprotective;  
KW nephrotropic; antithyroid; dermatological; immunomodulator;  
KW cell proliferation disorder; cell differentiation disorder;  
KW brain disorder; platelet disorder; breast disorder; colon disorder;  
KW kidney disorder; renal disorder; lung disorder; ovarian disorder;  
KW prostate disorder; cervical disorder; spleen disorder; thymus disorder;  
KW thyroid disorder; testes disorder; haematopoietic disorder;  
KW pancreatic disorder; skeletal muscle disorder; skin disorder;  
KW dermal disorder; bone metabolism disorder; immune disorder;  
KW inflammatory disorder; cardiovascular disorder;  
KW endobial cell disorder; liver disorder; viral disease; pain disorder;  
KW metabolic disorder; neurological disorder;  
KW central nervous system disorder; erythroid disorder;  
KW blood vessel disorder; angioleptic disorder; cancer; heart failure;  
KW hypertension; angina; osteoarthritis; rheumatoid arthritis;  
KW multiple sclerosis; Crohn's disease; psoriasis; asthma;

KW cell proliferation; cell differentiation; cell growth; cell division;  
KW human; gene; ss.  
XX OS Homo sapiens.  
XX PN US2004058355-A1.  
XX PD 25-MAR-2004.  
XX PF 25-APR-2003; 2003US-00423543.  
XX PR 30-SEP-1998; 98US-00163821.  
PR 27-JAN-1999; 99US-0117580P.  
PR 25-MAR-1999; 99US-00276400.  
PR 30-JUL-1999; 99US-00365162.  
PR 09-SEP-1999; 99US-00392189.  
PR 05-OCT-1999; 99US-00412210.  
PR 23-NOV-1999; 99US-00448076.  
PR 28-FEB-2000; 2000US-0186061P.  
PR 28-APR-2000; 2000US-0206888P.  
PR 19-MAY-2000; 2000US-0205447P.  
PR 30-JUN-2000; 2000US-00608921.  
PR 31-JUL-2000; 2000US-0221925P.  
PR 25-SEP-2000; 2000US-0234922P.  
PR 25-SEP-2000; 2000US-0235035P.  
PR 08-NOV-2000; 2000US-0246669P.  
PR 09-NOV-2000; 2000US-00711216.  
PR 14-NOV-2000; 2000US-0248325P.  
PR 15-NOV-2000; 2000US-0248993P.  
PR 22-DEC-2000; 2000US-0257511P.  
PR 05-JAN-2001; 2001US-0260166P.  
PR 28-FEB-2001; 2001US-00797039.  
PR 27-APR-2001; 2001US-00845044.  
PR 20-JUL-2001; 2001US-00909743.  
PR 31-JUL-2001; 2001US-00920346.  
PR 13-AUG-2001; 2001US-00928531.  
PR 14-AUG-2001; 2001US-00929218.  
PR 15-AUG-2001; 2001US-0312539P.  
PR 25-SEP-2001; 2001US-00963159.  
PR 08-NOV-2001; 2001US-00080816.  
PR 13-NOV-2001; 2001US-00012055.  
PR 15-NOV-2001; 2001US-00003690.  
PR 30-JAN-2002; 2002US-00060763.  
PR 25-MAR-2002; 2002US-00105989.  
PR 12-APR-2002; 2002US-00121911.  
PR 12-AUG-2002; 2002US-00217168.  
PR 22-OCT-2002; 2002US-00278036.  
PR 02-JAN-2003; 2003US-00336489.  
PR 03-JAN-2003; 2003US-00336153.  
XX PA (MILL-) MILLENNIUM PHARM INC.  
XX PI Kapeller-Libermann R, Hunter JJ, Meyers RE, Rudolph-Owen LA;  
PI Curtis RAJ, Olandt PJ, Tsai F, Galvin KM, Chun W, Williamson MJ;  
PI Silos-Santiago I, Bandaru R;  
XX WPI: 2004-268788/25.  
XX P-PSDB: ADL14161.  
XX New 21910, 56634, 55053, 2504, 15977, 14760, 25501, 17903, 3700, 21529,  
XX 26176, 26343, 56638, 18610, 33217, 21967, h1983, m1983, 38555 or 593  
XX nucleic acid molecules and proteins, useful for treating, e.g. cancer,  
XX heart failure and angina.  
XX Claim 1; SEQ ID NO 45; 139pp; English.  
XX The invention describes an isolated 21910, 56634, 55053, 2504, 15977,  
XX 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217,  
XX 21967, h1983, m1983, 38555 or 593 nucleic acid molecule (I) comprising  
XX any one of 40 nucleotide sequences (I). The nucleic acid molecules and  
XX polypeptides are useful for diagnosing and treating a subject having a  
XX disorder, or a subject at risk of developing a disorder, which is  
XX associated with aberrant 21910, 56634, 55053, 2504, 15977, 14760, 25501,



Db 1766 GAAGCTGCCTGAGCGCTGGCGGACGAGTCTTTTGGGGACAGCTGCTTTTCCCTGACAG 1825  
 QY 1835 ACTGCCAAGAGGTGACTGCGAGCTTACAGACAGCCCTAGGATCTGCTCAAAGCTCAGC 1893  
 Db 1926 ACTGCCAGGAGGTGACAGCGACCTTACCAGCAGGCACTGAGGGTCTGCTCAAAGCTCACC 1884  
 RESULT 4  
 ABK14000  
 ID ABK14000 standard; cDNA; 3353 BP.  
 AC  
 XX ABK14000;  
 XX  
 DT 02-JUL-2002 (first entry)  
 XX  
 DE cDNA encoding human protein kinase 3700.  
 XX  
 KW Human; protein kinase 3700; PK; protein phosphorylation; tumorigenesis;  
 KW cell signalling; mitogenesis; gene transcription; angiogenesis; sarcoma;  
 KW tissue repair; tissue regeneration; atherosclerosis; blood-brain barrier;  
 KW cell proliferation disorder; cell differentiation disorder; carcinoma;  
 KW haematopoietic neoplastic disorder; metastatic disorder; leukaemia;  
 KW cytoskeletal; antiatherosclerotic; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 157..2043  
 FT /\*tag= a  
 FT /product= "Protein kinase 3700"  
 FT /note= "The coding region (not including the terminator  
 FT codon) is specifically claimed in claims 7 and 28"  
 XX  
 PN WO200224921-A2.  
 XX  
 XX 28-MAR-2002.  
 XX  
 XX 25-SEP-2001; 2001WO-US030115.  
 XX  
 XX 25-SEP-2000; 2000US-0234922P.  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX  
 XX Curtis RAJ, Galvin KM;  
 XX  
 XX WPI; 2002-352007/38.  
 DR P-PSDB; AAU79652.  
 XX  
 XX Use of modulators of activity of 3700 protein for making medicament for  
 PT e.g., modulating protein phosphorylation or cell signaling, or for  
 PT treating or preventing cellular proliferative and/or differentiative  
 PT disorders.  
 XX  
 XX Claim 28; Fig 1; 115pp; English.  
 PS  
 XX The present invention relates to the isolation of a novel human protein  
 CC kinase designated 3700, and the polynucleotide sequence encoding it. The  
 CC invention also describes the use of a modulator of the activity of  
 CC protein kinase (PK) 3700 for making a medicament or pharmaceutical  
 CC composition for modulating the ability of a cell to phosphorylate an  
 CC amino acid residue of a substrate protein. Modulators of protein kinase  
 CC 3700 activity are useful for modulating protein phosphorylation, cell  
 CC signalling, tumorigenesis, mitogenesis, transcription of a gene,  
 CC angiogenesis, tissue repair, tissue regeneration, establishment or  
 CC progression of atherosclerosis, and signalling across the blood-brain  
 CC barrier. The polynucleotide and polypeptide molecules for protein kinase  
 CC 3700 may be used as diagnostic targets and therapeutic agents for  
 CC prognosticating, diagnosing, preventing, inhibiting, alleviating, or  
 CC curing PK-related disorders and cellular proliferative and/or  
 CC differentiative disorders (e.g. haematopoietic neoplastic disorders,  
 CC carcinoma, sarcoma, metastatic disorders or leukaemia). The 3700  
 CC polynucleotide sequence can be used to express protein kinase 3700, to  
 CC detect a genetic alteration in a 3700 gene, in chromosome mapping, for

CC tissue typing, in forensic biology, and as surrogate markers. The present  
 CC sequence encodes human protein kinase 3700  
 XX  
 SQ Sequence 3353 BP; 754 A; 1022 C; 879 G; 695 T; 0 U; 3 Other;  
 Query Match 71.9%; Score 1360.2; DB 6; Length 3353;  
 Best Local Similarity 84.2%; Pred. No. 0;  
 Matches 1565; Conservative 0; Mismatches 273; Indels 21; Gaps 2;  
 QY 47 CTCCGCCCTTGGCTCGGAGAGCGCCCGCGCTGGCGGACGGGCTCATCAAGTCGCGTA 106  
 Db 191 CCACTCCCTCGGCGCGCAGAGCTAGCCCGCGCTGGCGGAAGGCTCATCAAGTCGCGCA 250  
 QY 107 AACCTCTGATGAAGAAGCAGGCGGTGAAGCGGCACCATCAAAACACAACTCTGGGCACC 166  
 Db 251 AGCCCCCTAATGAAGAAGCAGGCGGTGAAGCGGCACCAACCAAGCAACCTCTGGGCACC 310  
 QY 167 GCTACAGATTCTCTGGAGAGCGCTGGGCAAGGGACCTTACGGGAAGGTGAAGAAGCAGCAG 226  
 Db 311 GCTACAGATTCTCTGGAGAGCCCTGGGCAAGGACACTTACGGGAAGGTGAAGAAGCGCGG 370  
 QY 227 AGAGCTCGGGGCGCTCTGGTGGCCATCAAGTCCATCAGGAAGACAAAATCAAAAGATGAGC 286  
 Db 371 AGAGCTCGGGGCGCTCTGGTGGCCATCAAGTCAATCCGAAGGACAAAATCAAAAGATGAGC 430  
 QY 287 AGGATCTGTGCACATACGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAACA 346  
 Db 431 AAGATCTGATGCACATACGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAACA 490  
 QY 347 TCATTGCCATCCATGAAGTGTGAGAAATAGCAGCAAGATTGTGATGTCATGGAGTATG 406  
 Db 491 TCATTGCCATCCATGAAGTGTGAGAAATAGCAGCAAGATTGTGATGTCATGGAGTATG 550  
 QY 407 CCAGCCGAGCGATCTGTATGATTATCATCAGTGAGCGGCGACCGCTGAGTGAGCGGAGCG 466  
 Db 551 CCAGCCGAGCGATCTGTATGATTATCATCAGTGAGCGGCGACCGCTGAGTGAGCGGAGCG 610  
 QY 467 CCAGGCGATTCTTCCGACAGATGCTCTGCCCTGCACTACTGTCACCAAGGAGCGGATCG 526  
 Db 611 CTAGGCGATTCTTCCGCGAGATGCTCTGCCCTGCACTACTGTCACCAAGGAGCGGATCG 670  
 QY 527 TTCACCGAGATCTCAAGCTGGAAAACATCTCTCTAGATGCCAATGGAACATCAAGATTG 586  
 Db 671 TCACCGAGATCTCAAGCTGGAAAACATCTCTCTAGATGCCAATGGAATATCAAGATTG 730  
 QY 587 CTGACTTTGGCTCTTCCAACTGTATACCAAAAGGCAAGTTCTTCCAGACGTTCTGTGGGA 646  
 Db 731 CTGACTTTGGCTCTTCCAACTGTATACCAAAAGGCAAGTTCTTCCAGACATTTCTGTGGGA 790  
 QY 647 GCGCTCTTACGCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCGGAGAGTGG 706  
 Db 791 GCGCTCTTATGCTCGCGAGAGATTGTCAATGGGAAGCCCTTACACAGGCGGAGAGTGG 850  
 QY 707 ACAGCTGGTCTCTGGGCGTTCTCTCTATACATCTCTGTCATGCGCACCACATGCCCCCTTTCACG 766  
 Db 851 ACAGCTGGTCTCTGGGCGTTCTCTCTATACATCTCTGTCATGCGCACCACATGCCCCCTTTCATG 910  
 QY 767 GGCAGGATCATAAACACTGGTGAACAAATCAGTAACCGGGGTTTACCGTGAGCGCGCCA 826  
 Db 911 GGCATGACCAATAGATCTTAGTGAACACAGATCAGCAACCGGGGCTTACCGGAGGACCACTA 970  
 QY 827 AGCGGTCCGATGCTGTGGCGTGTATGTCGGTGGTGTATGTTGATGGTGAACCCACCGCTCGGG 886  
 Db 971 AACCTCTGATGCTGTGGCGTGTATGTCGGTGGTGTATGTTGATGGTGAACCCACCGCTCGGG 1030  
 QY 887 CCACACTGGAGGATGTAGCCAGTCAATGTGGTGGTCACTGGGGGTTTACACACCGGAGTCG 946  
 Db 1031 CCACCTTGGAGGATGTGGCGAGTCACTGTGGTCACTGGGGCTACCGGAGTGG 1090  
 QY 947 GGGAAACAGGAGCCCTGCGTGGAGGTTGGGACCCCTAGTGGTGTGATTTGGCGGGGCTCCA 1006  
 Db 1091 GAGAGCAGGAGGCTCCGCATGAGGGTGGGCGCCCTGGCAGTGAATCTGCTGCCCGGCTCCA 1150









Db 1682 CCACAGAGCCTTAGAGCTCGCGCCGCCACACCTTCGGCTCCCTGGATGAATCGGCC 1741  
 QY 1607 CCTCCCATCTGACGCGCCGAGCCGCTCAGGGGCTGTGAGTGAGGACAGCATCC 1666  
 Db 1742 CACTCGCCCTCGCCGCGCCGAGCCGCTCAGGGGCTGTGAGGAGGACAGCATCC 1801  
 QY 1667 TGTCTCCGAGTCTTTGACCAATTTGACTTGGCTGAACGCTTTCCCGAAACCCCACTGA 1726  
 Db 1802 TGTCTCTGAGTCTTTGACCAAGCTGACCTTGGCTGAACGCTTCCAGAGCCCACTGC 1861  
 QY 1727 GGGCTGTGTCTGTGGACACCTGAGGGGCTTGGAGAGCTTCCCTCAGAG----- 1780  
 Db 1862 GGGCTGTGTCTGTGGACACCTGAGGGGCTTGGAGAGCTTCCCTCAGAGGGCCCTG 1921  
 QY 1781 -----GTCTGAAGCGATGTGGCAGGAATCTTTGGGGGATAGTCTTTCTCTGACAG 1834  
 Db 1922 GAAGCTGCTTGAAGCGCTGGCGGAGATGCTTTGGGGGAGAGTCTTTCTCTGACAG 1981  
 QY 1835 ACTGCCAAGAGTGAATGACCTGACGCTTACAGACAGCCCTTAGGAATCTGCTCAAGCTCAGC 1893  
 Db 1982 ACTGCCAAGAGTGAATGACGCTTACGACAGGACTGAGGGTCTGCTCAAGAGCTCACC 2040

## RESULT 6

AAD26459

ID AAD26459 standard; cDNA; 3360 BP.

AC AAD26459;

XX 26-MAR-2002 (first entry)

XX Human kinase PKIN-12 cDNA.

XX Human; kinase; PKIN-12; cancer; leukaemia; adenocarcinoma; osteoporosis;

XX Immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;  
 XX Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;  
 XX allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;  
 XX autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis;  
 XX Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;  
 XX rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;  
 XX hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;  
 XX cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;  
 XX congestive heart failure; ischaemic heart disease; lung tumour; gout;  
 XX fatty liver; Niemann-Pick's disease; gene therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FH 181..1965

FT CDS /\*tag= a

FT /\*product= "Human PKIN-12 protein"

XX W0200196547-A2.

XX 20-DEC-2001.

XX 14-JUN-2001; 2001WO-US019444.

XX 15-JUN-2000; 2000US-0212073P.

XX 23-JUN-2000; 2000US-0213467P.

XX 30-JUN-2000; 2000US-0215651P.

XX 07-JUL-2000; 2000US-0216605P.

XX 13-JUL-2000; 2000US-0218372P.

XX 25-AUG-2000; 2000US-0228056P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;

XX Gandhi AR, Tribouley CM, Walia NK, Yao MG, Lu DM, Greenwald SR;

XX Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang YT;

XX Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;

XX Lo TP, Khan F, Recipon SA, Azimzai Y, Policky JL, Ding L;

PI

XX

DR

DR

XX

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Sequence 3360 BP; 729 A; 1022 C; 882 G; 727 T; 0 U; 0 Other;

Query Match

Best Local Similarity 71.7%; Score 1357.6; DB 6; Length 3360;

Matches 1569; Conservative 0; Mismatches 284; Indels 21; Gaps 2;

QY 32 GCAGCGCTCCCTCGGCTCCGCTCGGCTCGGCTCGGAGAGCGCCGCGCTCGGCGAGCGGC 91

Db 98 GCAGCGCTCCGCGCCCACTCCCTCGGCGAGAGCTAGCCGCGCTCGGCGAGCGGC 157

QY 92 TCATCAAGTCCCTAAACCTCTGATGAAGAGAGCGCGGTGAAGCGGCACCATCACAAC 151

Db 158 TGATCAAGTCCGCCCAAGCCCTAATGAAGAAGAGCGCGGTGAAGCGGCACCATCACAAC 217

QY 152 ACAACCTCGGCGACCCGCTAGAGTTCCTGGAGAGCGCTGGGCAAGGCGACCTACGGGAGG 211

Db 218 ACAACCTCGGCGACCCGCTAGAGTTCCTGGAGAGCGCTGGGCAAGGCGACCTACGGGAGG 277

QY 212 TGAAGAAGGCGAGAGAGCTCGGCGCTCTGGTGCCCATCAAGTCCATCAGGAAGACA 271

Db 278 TGAAGAAGGCGCGGAGAGCTCGGCGCTCTGGTGCCCATCAAGTCCATCAGGAAGACA 337

QY 272 AATCAAGATGAGCAGGATCTCTGCACATACGAGGAGGAGTTGAGATCATGTCTTCAC 331

Db 338 AATCAAGATGAGCAGGATCTCTGCACATACGAGGAGGAGTTGAGATCATGTCTTCAC 397

QY 332 TCAACCAACCCCAACATCATTGGCCATCCATGAGTGTGAGAGTATGAGCAAGATTCTGA 391

Db 398 TCAACCAACCCCAACATCATTGGCCATCCATGAGTGTGAGAGTATGAGCAAGATTCTGA 457

QY 392 TTGTCTATGAGTATGCCAGCGGAGCGATCTGTATGATTACATCAGTACGAGCGGCACG 451

Db 458 TCGTCTATGAGTATGCCAGCGGAGCGATCTGTATGATTACATCAGTACGAGCGGCACG 517

QY 452 TGAGTACGCGGAGCGCAGGATCTTCTTCGACAGATCTGTCTGCTGCTGCTGCTGCTGCTG 511

Db 518 TCAGTACGCGGAGCGGAGGATCTTCTTCGCGCAGATCTGTCTGCTGCTGCTGCTGCTGCTG 577

Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;

WPI; 2002-050207/12.

P-PSDB; AAE16266.

New polypeptides, useful for diagnosing, treating or preventing disorders

of growth and development, cardiovascular and lipid, and diseases such as

cancer, comprise human kinase polypeptides.

Claim 5; Page 184; 197pp; English.

The invention relates to human kinase PKIN proteins and their

corresponding cDNAs. A composition containing PKIN agonist is useful for

treating a disease or condition associated with decreased expression of

PKIN and a composition comprising PKIN antagonist is useful for treating

a disease or condition associated with overexpression of PKIN. The

disorders include cancer (leukaemia, adenocarcinoma, lymphoma, melanoma,

myeloma, sarcoma, teratocarcinoma, Hodgkin's disease); immune disorder

(Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,

atherosclerosis, anaemia, allergies, adult respiratory distress syndrome,

autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes

mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,

osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,

rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,

bacterial, parasitic, fungal, viral, protozoal and helminthic infections)

growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,

Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); cardio

vascular disease (arteriovenous fistula, hypertension, vasculitis,

aneurysms, congestive heart failure, angina pectoris, myocarditis,

ischaemic heart disease, chronic bronchitis, lung tumours); lipid

disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,

hypocholesterolaemia, obesity). PKIN DNA is useful for assessing toxicity

of a test compound and in gene therapy. The present sequence is human

PKIN-12 cDNA

Qy	512	ACGAGAACGGGATCGTTTCAACCGAGATCTCAAGCTGGAAAAATCCTTCTAGATGCCAATG	571
Db	578	ATCAGAAACAGAGTTGTCTCACCGAGATCTCAAGCTGGAGAAATCCTCTTGGGTGCCAATG	637
Qy	572	GAAACATCAAGATTGCTGACTTTGGCCCTCTCCAACTCTGATACCAAAAGGCNAGTTCTCTCC	631
Db	638	GGAATATCAAGATTGCTGACTTTGGCCCTCTCCAACTCTACCATCAAGAGCAAGTTCTCTGC	697
Qy	632	AGACGTTCTGTGGAGGCCCTCTCTACGCTCCGCTGAGATAGTCAAACGGGAAGCCCTATG	691
Db	698	AGACATTTCTGTGGAGCCCCCTCTATGCCTGCCAGAGATTGTCAATGGGNAGCCCTACA	757
Qy	692	TGGCCACAGAGGTGGACAGCTGTGCTCTGTGGCGTTCTCCGTGATCATCTCGTGCATGGCA	751
Db	758	CAGGCCCAGAGGTGGACAGCTGTGCTTCCGTGGTGTCTCTCTACATCTGTGTGATGGCA	817
Qy	752	CCATGCCCTTTGACGGGCAGGATCATAAACACTGTGTGAAGCAATCAGTAAACGGGGCTT	811
Db	818	CCATGCCCTTTGATGGGCATGACCATAGATCTCTAGTGAATCCCTAGTGAACAGATCAGCAACGGGGCT	877
Qy	812	ACCGTGAGCCGCCCAAGCCGTCCGATGCTGTGGCCTGATCCGGTGGCTGTTAATGGTGA	871
Db	878	ACCGGAGACCACTTAACCTCTCATGCTGTGGCTGATCCGGTGGCTGTTAATGGTGA	937
Qy	872	ACCCACCCGCTCGGGCCACACTGAGAGATGTAGCCAGTCAATGTGTGGGTCAACTGGGGTT	931
Db	938	ACCCACCCGCGGGCCACCCCTGGAGGATGTGGCCAGTCACTGGTGGGTCAACTGGGGCT	997
Qy	932	ACACCACGGAGTCGGGGMACAGAAAGCCCTGCTGAGGGTGGGCACCTAGTGGTGACT	991
Db	998	ACGCCACCCGAGTGGGAGAGCAGAGAGCTCCGCATGAGGGTGGGCACCTCTGGCAGTACT	1057
Qy	992	TTGGCCGGGCTCCATGCGGCACTGTGTTCGTCTCTCGCGCCCTCTCTGGAGAAATG	1051
Db	1058	CTGCCCGGCTCCATGGCTGACTGTGGCTCGGCGTTCTCCGCGCCCTCTCTGGAGAAATG	1117
Qy	1052	GAGCCAAAGGTGTGAGCTTTCTTAAGCAGACAGTCGCCGGAGGTGGAGCACTGTACTCTG	1111
Db	1118	GGGCCAAAGGTGTGAGCTTTCTTAAGCAGCATGCACCTGGTGGGGAGACACACCCCTG	1177
Qy	1112	GGCTGAGCGGCACATTTCTTTAAGAGTCCCGAAAGGAGATGAATGGCTCAAAATC	1171
Db	1178	GCCTGGAGCCACAGATTCGCTCAAGAAATCCCGCAAGGAGAAATGAATGGCCCAAGTCTC	1237
Qy	1172	TGCAAGGTGACCGGCTGAGATACCTTTCTCGCCCTTGGCAAGCAGACGCTTAAGCTTC	1231
Db	1238	TCCACAGTGACCGGCTGTATGACTGCCCATCGCCCTGGCAGAGACAACTCAAGCTGC	1297
Qy	1232	CGAAAGCATTTCTCAAGAAAAGTCTCTACTCTCGTCAGGGGAGGTACAGAGGACCCCTC	1291
Db	1298	CAAAAGGCATTTCTCAAGAAAGGTGTGAGCTCTGCAGAAAGGGGTACAGGAGGACCCCTC	1357
Qy	1292	AGGAACTCAGACCGGTGCTGATTAATCTCAGGGCAGCTGTCTCTGTGTATTCCTCTGTCTC	1351
Db	1358	CGGAGCTCAGCCCAATCCCTCGAGCCACGGGCAAGGTGCCCC-----CCCTGTCTCC	1408
Qy	1352	CAAGGAAGGCATCTTTAAGAAATCTCGACAGCGTGAATCTGGTTACTCTCTCTCTCCAG	1411
Db	1409	CCAAAGAGGCATTTCTCAAGAAAGCCCGACAGCCGAGTCTGGCTACTCTCTCTCTCCCG	1468
Qy	1412	AGCCCAGCAGTCTGGGGAACTCTTAGACGCAGTGTATGTGTTGTGAGTGGGACCCCG	1471
Db	1469	AGCCAGTGAATCTGGGGAGCTCTTGGACGAGCGACGCTGTTGTGTGAGTGGGATCCCA	1528
Qy	1472	TGGAGCAGAGTCTCCAAGGCTTACGGGTCTCTCTCCACCGCAAGGGCAATTCCTAAAC	1531
Db	1529	AGGAGCAGAAAGCTCTCGCAAGCTTTCAGGGTGTCTCTCCATCGCAAAAGGCATCTCTAAAC	1588
Qy	1532	TCAATGCAAGTTCTCCGCGACACGCTTTAGAAAGCACTACCCCTAGCACCTTTGGCTCCC	1591
Db	1589	TCAATGCAAGTTCTCCCAAGACGCTTTGAGTCTCGGGGCCCCCACACCTTTGGGTCCC	1648
Qy	1592	TGGAACCAACTGGGCTCTCTCCATCTCTGCAGGCCCGGGCCAGCGCCCTCTAGGGGCTGTGA	1651

RESULT 7

AK94280  
RESOL

AAK94280  
ID AAK94280 standard: cDNA: 3395 BP.

AA  
AC AAK94280:

DT 06-NOV-2001 (first entry)

Human full-length cDNA, SEO ID NO: 2918.

Human: full length cDNA: cDNA synthesis: oligo-capping: ss:

OS Homo sapiens.

XX  
PN  
EP1130094-A2.

05-SEP-2001.

07-JUL-2000; 2000EP-00114089.

08-JUL-1999: 99JP-00194486.

PR II-UAN-2000; 2000JP-0018774;  
PR 02-MAY-2000; 2000JP-00183765;

PA (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, I

FI waxaannacshu n, sugiyama i, nayaat n, wajiima s, ucsaki i, wogga  
XX

DR P-PSDB; AAM933360.

PT 830 Primers useful f

XX

XX

CC clones. 830 cDNA molecule

CC been determined. Primers for synthesising the full length cD

length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a full length human cDNA of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO

XX	SQ	Sequence	3395 BP; 722 A; 1044 C; 894 G; 735 T; 0 U; 0 Other;
Query Match		71.7%;	Score 1357; DB 4; Length 3395;
Best Local Similarity		84.1%;	Pred. No. 0;
Matches 1563; Conservative		0;	Mismatches 275; Indels 21; Gaps 2;
Qy	47	CCTCGGCCCTCGCGAGAGCCCGCGCGCTGGCGACGGGCTCATCAAGTCGGCTA	106
Db	152	CCAATCCCTCGCCGCGAGCTAGCCCGCGCTGGCGAAGGCTGATCAAGTCGGCCA	211
Qy	107	AACCTCTGATGAAGAAGCAGCGGTGAAGCGGCACATCAAAACACAACTCGCGGAC	166
Db	212	AGCCCCCTAATGAAGAAGCAGCGGTGAAGCGGCACCAACAGACACAACTCGCGGAC	271
Qy	167	GCTACGAGTTCTCTGGAGACGCTGGCGAAGGSCACTACGGGAAGGCTGAAGAAGCACGAG	226
Db	272	GCTACGAGTTCTCTGGAGACCTGGCGAAGGSCACTACGGGAAGGCTGAAGAAGCGCGGG	331
Qy	227	AGAGCTCGGGCGCTCTGTGGCCATCAAGTCCATCAGGAAAGA CAAATCAAGATGAGC	286
Db	332	AGAGCTCGGGCGCTGTGTGGCCATCAAGTCAATCCGGAAGGACAAATCAAGATGAGC	391
Qy	287	AGGATCTGTGCACATACGAGGGGAGATTGAGATCATGTCTTACTCAACACCCCCACA	346
Db	392	AAGATCTGTGCACATACGAGGGGAGATTGAGATCATGTCTACTCAACACCCCTCACA	451
Qy	347	TCATTGCCATCATGAAGTGTTCAGAAATAGCAGCAAGATTGTGATTGTCTATCGAGTATG	406
Db	452	TCATTGCCATCCATGAAGTGTTCAGAAACAGCAGCAAGATCGTGATCGTATGGAGTATG	511
Qy	407	CCAGCCGAGGCGATCTGTATGATTACATCAGTGAAGCGGCCACGGCTGAGTCAGCGGAGC	466
Db	512	CCAGCCGGGCGACCTTTATGACTACATCAGCAGCGCGCAGCTCAGTGAAGCGGAG	571
Qy	467	CCAGGCATTTCTCCGACAGATCGTGTCTGCGCTGCACTACTGCGCACAGAACGGGATCG	526
Db	572	CTAGGCATTTCTCCGCGAGATCGTCTCTGCGGTGCATATTGCGCATCAGAACAGAGTTG	631
Qy	527	TTCAACCGAGATCTCAAGCTGGAACATCCTTCTAGATGCCAATGGAAACATCAAGATTG	586
Db	632	TCCACCGAGATCTCAAGCTGAGAACATCCTCTTGGATGCCAATGGGAATATCAAGATTG	691
Qy	587	CTGACTTTGGCCTCTCCAAACCTGTACACAAAGGCAAGTTCTCTCAGACGTTCTGTGGGA	646
Db	692	CTGACTTCGGCTCTCCAACTCTACCATCAAGGCAAGTTCTCTGCACATTTCTGTGGGA	751
Qy	647	GCCCTCTTACGCTTCGCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCACAGAGTGG	706
Db	752	GCCCCCTCTATGCTTCGCCACAGATTGTCAATGGGAAGCCCTACACAGGCGCCAGAGTGG	811
Qy	707	ACAGCTGGTCTCTGGGCGTTCTCTGTACATCCTGTGTGATGCGACCATGCGCTTTGACG	766
Db	812	ACAGCTGGTCCCTGGGTGTTCTCTCTACATCTGTGTGATGCGACCATGCGCTTTGATG	871
Qy	767	GGCAGGATCATAAAACA CTGGTGAAGCAAAATCAGTAAACGGGGCTTACCGTGAGCGGCCA	826
Db	872	GGCATGACCATAGATCCTAGTGAACAGATCAGACAGGGGCTTACCGGAGGCCACTTA	931
Qy	827	AGCGGTCCGATGCCCTGTGGCTGATTCGGTGGCTGTTTAATGGTGAACCCCAACCGCTCGG	886
Db	932	AAACCTCTGATGCCCTGTGGCTGATTCGGTGGCTGTTTAATGGTGAACCCCAACCGCTCGG	991
Qy	887	CCACATCGAGGATGTAGCAGTCAATGGTGGGTCAACTGGGGTTACACACCGGAGTCG	946
Db	992	CCACCTCGAGGATGTGGCAGTCACTGTGTGGGTCAACTGGGGTTACACACCGGAGTCG	1051
Qy	947	GGGAAACAGGAAGCCCTCGCTGAGGGTGGGCAACCTAGTGTGATCTTTGGCGGGCCCTCCA	1006
Db	1052	GAGAGCAGAGGCTCCGCATGAGGGTGGGCAACCTTGCAGTGAATCTGCGCGGCTTCCA	1111
Qy	1007	TGGCGGACTGGTTACGTGGTCTCTCGCGCCCCCTCTCTGGAGAAATGGAGGCAAGGTGTGCA	1066

XX EPI396543-A2.  
XX 10-MAR-2004.  
XX 07-JUL-2000; 2003EP-00025638.  
XX 08-JUL-1999; 99JP-00194486.  
XX 11-JAN-2000; 2000JP-00118774.  
XX 02-MAY-2000; 2000JP-00183865.  
XX 07-JUL-2000; 2000EP-00114089.  
XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX Oca T, Nishikawa T, Isogai T, Hayaashi K, Ishii S, Kawai Y;  
XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX WPI; 2004-204755/20.  
XX P-PSDB; ADL30886.  
XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full  
XX length human cDNAs.  
XX Example 1; SEQ ID NO 2918; 1340pp; English.  
XX This invention relates to a novel primers useful for synthesizing full  
XX length cDNA molecules that encode human proteins. Specifically, it refers  
XX to secretory or membrane proteins that are potential therapeutic agents/  
XX target molecules in the field of medicine, and in particular genes  
XX encoding proteins that are associated with signal transduction,  
XX glycoproteins and transcription. The present invention describes a method  
XX for efficiently cloning a full length human cDNA from both the 5' and 3'  
XX ends using the oligo-capping method. This polynucleotide sequence is a  
XX full length human cDNA clone of the invention.  
XX  
SQ Sequence 3395 BP; 722 A; 1044 C; 894 G; 735 T; 0 U; 0 Other;  
Query Match 71.7%; Score 1357; DB 12; Length 3395;  
Best Local Similarity 84.1%; Pred. No. 0;  
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;  
47 CCTCCGCCCTGGCTCGAGAGCGCCCGCGCTGGCGGAGGCTCATCAAGTCGCGTA 106  
152 CCACTCCCTCGCGCGCAGAGCTAGCCCGCGCTGGCGGAGGCTGATCAAGTCGCGCA 211  
107 AACCTCTGATGAAGACGCGCGTGAAGCGGACCATCAAAACACAACTCGCGGACC 166  
212 AGCCCTTAATGAAGACGCGCGTGAAGCGGACCATCAAAACACAACTCGCGGACC 271  
167 GCTACGAGTTCTCGAGACGCTGGCGCAAGGCGACCTACGGGAGGTGAAGACGAG 226  
272 GCTACGAGTTCTCGAGACGCTGGCGCAAGGCGACCTACGGGAGGTGAAGACGAG 331  
227 AGAGCTCGGGCGCTTGGTGCCATCAAGTCAATCAGGAAGACAAATCAAGATGAGC 286  
332 AGAGCTCGGGCGCTTGGTGCCATCAAGTCAATCAGGAAGACAAATCAAGATGAGC 391  
287 AGATCTGCTGCACATACGGGAGGAGATTGATGATCTTTCACTCAACCCACCCACA 346  
392 AAGATCTGATGCACATACGGGAGGAGATTGATGATCTTTCACTCAACCCACCCACA 451  
347 TCATTGCGCATCCATGAAGTGTTCAGAAATAGCAGCAAGATTGTGATGATGATG 406  
452 TCATTGCGCATCCATGAAGTGTTCAGAAATAGCAGCAAGATTGTGATGATGATG 511  
407 CCAGCGAGGCGATCTGTATGATTAATCATAGTAGCGGCCACCGCTGAGTAGCGGACG 466  
512 CCAGCGGCGGCGCTTATGACTACATCAGGAGCGGCGAGCTCAGTGAGCGCGAAG 571  
467 CCAGGCAATTTCTCCGACAGATCGTGTGCGCTGCACTACTGCCACAGAACCGGATCG 526  
572 CTAGGCAATTTCTCCGCGAGATCGTGTCTCGCGTGCATATTGCGCATCAGAACAGATG 631

QY TTCAACCGAGATCTCAAGCTGGAAAAACATCTTTAGATGCCAATGGAAAAACATCAAGATTG 586  
DB |||||  
QY TCCACCGAGATCTCAAGCTGGAGAACATCTCTTGGATGCCAATGGGAATATCAAGATTG 691  
DB |||||  
QY CTGACTTTGGCTCTCCAACTCTGATACCAAAAGGCAAGTTCTCTCCAGAGCTTCTGTGGGA 646  
DB |||||  
QY CTGACTTTGGCTCTCCAACTCTGATCAAAAGGCAAGTTCTCTCCAGAGCTTCTGTGGGA 751  
DB |||||  
QY GCGCTCTCTACGCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCGCCAGAGGTGG 706  
DB |||||  
QY GCGCTCTCTATGCTCGCTGAGATAGTCAATGGGAAGCCCTATACAGGCGCCAGAGGTGG 811  
DB |||||  
QY ACAGCTGGTCTCTGGGCGTTCTCTCTATACATCTCTGTCATGGCAACATCGCCCTTTGACG 766  
DB |||||  
QY ACAGCTGGTCTCTGGGCGTTCTCTCTATACATCTCTGTCATGGCAACATCGCCCTTTGATG 871  
DB |||||  
QY GGAGGATCATAAACACATGGTGAAGCAATAGTAACGGGGCTTACCGTGAGCGGCCCA 826  
DB |||||  
QY GGATGACCATAGATCTTAGTGAACAGATCAGCAACGGGGCTTACCGGGAGCCACTA 931  
DB |||||  
QY AGCGTCCGATGCTGTCGGCTGATCCGGTGGCTGTATGTGTGAACCCACCCCGCTGGG 886  
DB |||||  
QY AACCTCTGATGCTGTCGGCTGATCCGGTGGCTGTATGTGTGAACCCACCCCGCTGGG 991  
DB |||||  
QY CCACACTGGAGATGTAGCCAGTCAATTGGTGGTCAACTGGGGTTTACACACCGGAGTCG 946  
DB |||||  
QY CCACCTTGGAGATGTGGCCAGTCACTGGTGGTCAACTGGGGCTACGCCACCGAGTGG 1051  
DB |||||  
QY GGAAACAGAGACCTGCTGGTGGGTCGAGGTCGAGGTCGAGTGGTGGTGGTGGTGGTGGTGG 1006  
DB |||||  
QY GAGAGCAGGAGGCTCCGCTGAGGTCGAGGTCGAGGTCGAGTGGTGGTGGTGGTGGTGGTGG 1111  
DB |||||  
QY TGGCGGACTGTTTACGTCGCTCTCGCGCCCTCTCTGGAGATGAGGTCGAGGTCGAGTGGTGG 1066  
DB |||||  
QY TGGCTGCTGCTCGCGGCTTCTCTCGCGCCCTCTCTGGAGATGAGGTCGAGGTCGAGTGGTGG 1171  
DB |||||  
QY GCTTCTTCAAGCAGACGTCGCGGAGGTGGAAGCACTGTACCTGGGCTGGAGCGGCAAC 1126  
DB |||||  
QY GCTTCTTCAAGCAGACGTCGCGGAGGTGGAAGCACTGTACCTGGGCTGGAGCGGCAAC 1231  
DB |||||  
QY ATTCTCTTAAGAGTCCGAAAGGAGATGACATGCTCAAAATCTGCAAGGTGACCCGG 1186  
DB |||||  
QY ATTCTCTTAAGAGTCCGCAAGGAGATGACATGCTCAAAATCTGCAAGGTGACCCGG 1291  
DB |||||  
QY CTGAGGATACCTCTCTCGCCCTGGCAAGCAGACGCTTAAGCTTCCGAAAGGATTTCTCA 1246  
DB |||||  
QY CTGATGACATGCGCATCGCCCTGGCAGAGCACTCAAGCTGCCAAGGAGATTTCTCA 1351  
DB |||||  
QY AGAAAAGTCTCTTACCTCGTCAAGGAGGTACAGAGGACCTCTCAGGAATCTCAGACCGG 1306  
DB |||||  
QY AGAAGAGGTGTGAGCTCTGCAAGAGGAGGTACAGAGGACCTCTCGGAGCTCAGCCCAA 1411  
DB |||||  
QY TGCTGATCTCAGGCGACCTGTCCCTGCTGTATCCCTGCTCCCAAGGAGGAGTCC 1366  
DB |||||  
QY TCCCTCGAGCCCGAGGCGAGCTGCCCC-----CCCTGCTCCCAAGGAGGAGTCC 1462  
DB |||||  
QY TTAAGAGTCTCGACACGCTGAATCTGTTACTTCTCTCCAGAGCCCGAGGAGTCTG 1426  
DB |||||  
QY TCAAGAGCCCGACAGCGGAGTCTGGTACTTCTCTCCGAGCCCGAGTGAATCTG 1522  
DB |||||  
QY GGAAGCTTTAGACGCGAGT 1486  
DB |||||  
QY GAGAGCTTTGGAGCGAGGCGAGT 1592  
DB |||||  
QY CAGAGCTTCAGGCTCTCTCCAGGAGGAGTCTTCAAACTCAATGGAAGTCT 1546  
DB |||||  
QY CGAAGCTTCAGGCTCTCTCCAGGAGGAGTCTTCAAACTCAATGGAAGTCT 1642  
DB |||||  
QY CCGGCAAGCTTTAGAGGCACTACCCCTAGACCTTTGGCTCCCTGGGCACTGCGCT 1606  
DB |||||  
QY CCGAGAGCTTTGAGGCTCGCGGCCCGCCACCTTGGCTCCCTGGATGACTCGCC 1702  
DB |||||  
QY CTTCCCATCTGCGAGCCCGGCCCGCCCTCAGGGGCTGTGAGTGAAGACAGCATCC 1666

Db 1703 CACCTCCGCCCTTGGCCCGCCAGCCCTCAGGGGCTGTAGCGAGACAGCATCC 1762  
 QY 1667 TGTCTCCGAGTCTTTTGACCAATTGGACTTTGGCTGAACGCTTTCGCCGAAACCCCACTGA 1726  
 Db 1763 TGTCTCTGAGTCTTTTGACCAAGTGGACTTTGGCTGAACGCTTTCGCCGAAACCCCACTGC 1822  
 QY 1727 GGGGCTGTGTCTGTGGCAACCTGAGGGGGCTTGGAGAGCCCTCCCTCAGAAG----- 1780  
 Db 1823 GGGGCTGTGTCTGTGGCAACCTCAGGGGGCTTGGAGAGCCCTCCCTCAGAAG----- 1780  
 QY 1781 -----GTCCTGACCGATGGTGGCAGGAATCTTGGGGGATAGTCTTTTCTTGACAG 1834  
 Db 1883 GAAGCTCCCTGAGCGCTGGCGGAGGATCTTTGGGGGACAGCTGCTTTTCCCTGACAG 1942  
 QY 1835 ACTGCCAAGAGGTGACTGTCAGCCTTACAGACAGCCCTAGGAATCTGCTCAAAGCTCAGC 1893  
 Db 1943 ACTGCCAAGAGGTGACAGCGACCTTACAGACAGGACTGAGGGTCTGCTCAAAGCTCACC 2001  
 RESULT 9  
 ADJ96554  
 ID ADJ96554 standard; DNA; 3463 BP.  
 XX ADJ96554;  
 AC  
 XX  
 DT  
 XX  
 06-MAY-2004 (first entry)  
 DE Human calcium/calmodulin-dependent protein kinase NuaK2 DNA SeqID 11.  
 XX  
 KW gene; ds; kinase; human; SNP; single nucleotide polymorphism;  
 KW tyrosine protein kinase; serine/threonine protein kinase; PTK; STK;  
 KW gene therapy; cancer; immune-related disease; cardiovascular disease;  
 KW brain; neuronal associated disease; metabolic; inflammatory disorder;  
 KW cytosstatic; neuroprotective; immunomodulator; antiinflammatory;  
 KW calcium/calmodulin-dependent protein kinase; NuaK2.  
 XX  
 OS Homo sapiens.  
 OS 68.  
 XX  
 FH Key Location/Qualifiers  
 FT variation replace(1670,c)  
 FT /\*tag= a  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(1727,a)  
 FT /\*tag= b  
 FT /standard\_name= "Single nucleotide polymorphism"  
 XX  
 PN W02004006838-A2.  
 XX  
 PD 22-JAN-2004.  
 XX  
 PF 15-JUL-2003; 2003WO-US021730.  
 XX  
 PR 15-JUL-2002; 2002US-0395632P.  
 XX  
 PA (SUGEN-) SUGEN INC.  
 XX  
 PI Whyte D, Manning G, Caenepeel S;  
 XX  
 DR WPI; 2004-122753/12.  
 DR P-PSDB; ADJ96620.  
 XX  
 PT New nucleic acid molecule encoding a kinase polypeptide, useful for  
 PT preparing a composition for treating diseases or disorders, e.g., cancer,  
 PT or neurological, immunological or inflammatory disorders.  
 XX  
 PS Example 1; SEQ ID NO 11; 366pp; English.  
 XX  
 CC This invention relates to a novel isolated, enriched or purified nucleic  
 CC acid molecule that encodes a kinase polypeptide. Specifically, it relates  
 CC to human tyrosine and serine/threonine protein kinases (PTK's and STK's),  
 CC as well as protein kinase-like enzymes. The present invention describes

CC screening methods to identify agonists, antagonists and antibodies that  
 CC can be used to modulate the activity or function of the mammalian kinase  
 CC enzymes. As such, these compositions can be used for gene therapy  
 CC purposes to treat diseases or disorders including cancer, immune-related  
 CC diseases, cardiovascular disease, brain or neuronal associated disease,  
 CC metabolic and inflammatory disorders. Accordingly, they exhibit  
 CC cytostatic, neuroprotective, immunomodulator and antiinflammatory  
 CC activities. This polynucleotide sequence is a human kinase DNA sequence  
 CC of the invention.  
 XX  
 SQ Sequence 3463 BP; 732 A; 1057 C; 919 G; 755 T; 0 U; 0 Other;  
 Query Match 71.7%; Score 1357; DB 12; Length 3463;  
 Best Local Similarity 84.1%; Pred. No. 0;  
 Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;  
 QY 47 CTTCCGCCCTGGGCTCGAGAGCGCCCGCGCTGGCGAGCGGCTCATCAAGTCGGCTA 106  
 Db 223 CCACCTCCCTCGCGCCGAGAGCTAGCCCGCGCTGGCGGAAGGCTGATCAAGTCGCCCA 282  
 QY 107 AACCTCTGATGAAGAAGCAGCGGCTGAAGCGGCACCATCAAAACACAACTCTCGGCACC 166  
 Db 283 AGCCCTTATGAAGAAGCAGCGGCTGAAGCGGCACCATCAAAACACAACTCTCGGCACC 342  
 QY 167 GCTACGAGTTCTCTGGAGACGCTGGGCAAGGCGCACCTACGGGAAGGTGAAGAAGCGACG 226  
 Db 343 GCTACGAGTTCTCTGGAGACCTCTGGCAAGGCGCACCTACGGGAAGGTGAAGAAGCGCGG 402  
 QY 227 AGAGCTCGGGGCTCTGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGC 286  
 Db 403 AGAGCTCGGGGCGCTGTGGCCATCAAGTCAATTCGGGAAGGACAAATCAAGATGAGC 462  
 QY 287 AGGATCTGTGCACATACCGAGGAGATTGAGATCATGCTTCACTCAACCAACCCCA 346  
 Db 463 AAGATCTGTGCACATACCGAGGAGATTGAGATCATGCTTCACTCAACCAACCCCA 522  
 QY 347 TCATTGCCATCCATGAAGTGTTTGAGATAGCAGCAAGATTGTGATTGTCTCATGAGTATG 406  
 Db 523 TCATTGCCATCCATGAAGTGTTTGAGACAGCAGCAAGATCGTGTATGCTCATGAGTATG 582  
 QY 407 CAGCGCGAGCGATCTGTATGATTATCATAGTGAAGCGCCACCGCTGAGTGAGCGGACG 466  
 Db 583 CAGCGCGAGCGATCTGTATGATTATCATAGTGAAGCGCCACCGCTGAGTGAGCGGACG 642  
 QY 467 CAGGCGATTTCTTCGACAGATCGTCTGCGCTGCACTACTGCCACCAAGCGGATCG 526  
 Db 643 CTAGGATTTCTTCCGCGAGATCGTCTGCGCGTGCATCTATGCCATCAGAACAGAGTTG 702  
 QY 527 TTCACCGAGATCTCAAGCTGGAAACATCCTTCTAGATGCCAATGGAAACATCAAGATTG 586  
 Db 703 TCACCGAGATCTCAAGCTGGAAACATCCTTCTGGATGCCAATGGAAATATCAAGATTG 762  
 QY 587 CTGACTTTGGGCTCTCAACCTGTATACCAAGGCAAGTTCTCCAGACCTTCTGTGGGA 646  
 Db 763 CTGACTTCGGGCTCTCAACCTGTATACCAAGGCAAGTTCTCCAGACATCTCTGTGGGA 822  
 QY 647 GCCTCTCTACGCTCGCCTGAGATAGTCAACCGGGAAGCCCTATGTGGGCGCAGAGGTGG 706  
 Db 823 GCGGCTCTATGCTCGCAGAGATTGTCAATGGGAAGCCCTACACAGGCGCCAGAGGTGG 882  
 QY 707 ACAGCTGGTCTCTGGGCGTTTCTCTGTATACCTCTGTGATGCACCATCCCTTTTGACG 766  
 Db 883 ACAGCTGGTCTCTGGGCGTTTCTCTGTATACCTCTGTGATGCACCATCCCTTTTGATG 942  
 QY 767 GGCGAGATCATAAAACTGTGGTGAAGCAATCAAGTAAACGGGGCTTACCGTGAAGCGGCCCA 826  
 Db 943 GGCGATGACCAATAGATCTAGTGAAGCAATCAAGTAAACGGGGCTTACCGGAGCGCCCTA 1002  
 QY 827 AGCGGCTCCGATGCTGTGGGCTGATCCGGTGGCTGTTTATGGTGAACCCCAACCGCTGGG 886  
 Db 1003 AACCTCTGATGCTGTGGGCTGATCCGGTGGCTGTTTATGGTGAACCCCAACCGCTGGG 1062  
 QY 887 CCACACTGGAGGATGTAGCGAGTCAATTGGTGGGTCAATTGGGTTTACACCCAGGAGTCG 946

Db 1063 CCACCTGGAGGATGTGGCCAGTCACTGGTGGTCAACTGGGGCTAGCCACCCGAGTGG 1122  
Qy 947 GGGACAGGAAGCCCTGGGTGAGGGTGGGACCCCTAGTGTGATCTTTGGCCGGGCTTCCA 1006  
Db 1123 GAGAGCAGGAGGCTCCGATGAGGGTGGGACCCCTGGCAGTGAATCTGCCCGGCTTCCA 1182  
Qy 1007 TGGCGGACTGGTTACGTCTCCTCGGCCCTCCTGGGAGAAATGGAGCACAAGGTGGCA 1066  
Db 1183 TGGCTGACTGGCTCGGGTTCCTCCGCCCTCCTGGGAAATGGGCGCAAGGTGGCA 1242  
Qy 1067 GCTTCTTCAAGCAGCAGTGGCGGAGGTGGAAGCACTGTACCTGGGCTGGAGCGGCAAC 1126  
Db 1243 GCTTCTTCAAGCAGCAGTGCCTGGTGGGGAAGCACCACCCCTGGCTGGAGCGCCAGC 1302  
Qy 1127 ATTCTCTTAAGAAGTCCGAAAGAGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGG 1186  
Db 1303 ATTCTCTTAAGAAGTCCGAAAGAGAGAAATGACATGGGCTCAAAATCTGCAAGGTGACCCGG 1362  
Qy 1187 CTGAGGATACCTTCTCGCCCTGGCAAGAGCAGCCTTAAGCTTCCGAAAGGCAATCTCA 1246  
Db 1363 CTGATGACATGGCCCTGGCAAGAGCAACCTCAAGCTGCCAAAGGCAATCTCA 1422  
Qy 1247 AGAAAAAGTCTTACCTCTCGTCAAGGAGGTACAGGAGGACCTCTCAGGAACTCAGACCCG 1306  
Db 1423 AGAAGAGGTGTGAGCTCTGCAAGAGGGGTACAGGAGGACCTCTCGGAGCTCAGCCCAA 1482  
Qy 1307 TGCTGTATCTCAGGCGAGCTGTCTCTGCTGTATCTCTGCTCCCAAGAAAGGCAATCC 1366  
Db 1483 TCCCTGCGAGCCAGGCGAGGCTGCC-----CCCTGCTCCCAAGAAAGGCAATCC 1533  
Qy 1367 TTAAGAGTCTCGACAGCGTGAATCTGTTACTACTCTCTCCAGAGCCCGAGGCTG 1426  
Db 1534 TCAAGAGCCCGACAGCGAGTCTGGCTACTACTCTCTCCGAGCCAGTGAATCTG 1593  
Qy 1427 GGAAGCTCTTAGAGCGCAGTGTGTTGTGAGTGGGAGCCCGCTGGAGCAGAAAGTCTC 1486  
Db 1594 GGGAGCTCTTGGAGCGAGCGAGCTGTTGTGAGTGGGATCCCAAGGAGCAGAGCCTC 1653  
Qy 1487 CACAGGCTTCAAGGCTCTCTCTCAGCGAGGCAATCTCAAACTCAATGGCAAGTCT 1546  
Db 1654 CGCAAGCTTCAAGGCTCTCTCTCAGCGAGGCAATCTCAAACTCAATGGCAAGTCT 1713  
Qy 1547 CCCACAGCTTGAAGGCACTACCTCTAGCAGCTTGGCTCTCTGAGCAGCACTGAGCT 1606  
Db 1714 CCAGACAGCTTGGAGCTCGGCGCCCGCCACCTTGGCTCTCTGAGTGAATCTGCGCC 1773  
Qy 1607 CTTCCCACTCTGAGCGCGCCCGCCAGCCCTCAGGCGGCTGTGAGTGAGGACAGCATCC 1666  
Db 1774 CACCTCGCCCTTGGCGCGGCGAGCGAGCCCTCAGGCGGCTGTGAGTGAGGACAGCATCC 1833  
Qy 1667 TGTCTCTGAGTCTTTGACCAATGAGTCTGCTGAAGCTTCTCCGAAACCCCACTGA 1726  
Db 1834 TGTCTCTGAGTCTTTGACCAATGAGTCTGCTGAAGCTTCTCCGAAACCCCACTGC 1893  
Qy 1727 GGGCTGTGTGTCTGTGCAACCTGAGGGGGCTTGAAGAGCTCTCCCTCAGAG----- 1780  
Db 1894 GGGCTGTGTGTCTGTGCAACCTTGAAGAGCTTGAAGAGCTCTCCCTCAGAG----- 1953  
Qy 1781 -----GTCTGAAGCGATGTGGCAGGAATCTTTGGGGGATAGTGTCTTTCTGACAG 1834  
Db 1954 GAAGTCTGCTGAGCGCTGGCGGAGGATCTTTGGGGGAGAGTGTCTTTCTGACAG 2013  
Qy 1835 ACTGCCAAGAGGTGACTGAGCTTACAGACAGCCCTTAGGAATCTGTCTAAGCTGAGC 1893  
Db 2014 ACTGCCAAGAGGTGAGCAGCTTACCGAGGAGCTGAGGGTCTGTCTAAGAGCTCACC 2072

RESULT 10  
ABZ11333  
ID ABZ11333 standard; cDNA; 2043 BP.  
XX  
AC ABZ11333;

XX 20-JAN-2003 (first entry)  
DT Human polynucleotide SEQ ID NO 215.  
XX  
DE  
XX  
XX  
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;  
cell-proliferative disorder; neurodegenerative disease; bacterial;  
Parkinson's disease; Alzheimer's disease; autoimmune disease;  
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
arthritis; cytostatic; immunomodulator; neurotropic; neuroprotective;  
antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
haemostatic; vulnary; fungicide; antibacterial; virucide; protozoacide;  
antiarthritic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200270539-A2.  
XX  
PD 12-SEP-2002.  
XX  
PF 05-MAR-2002; 2002WO-US005095.  
XX  
PR 05-MAR-2001; 2001US-00799451.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;  
Xue AU, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
Wehrman T, Wang J, Wang D, Drmanac RT;  
XX  
PI WPI; 2002-759812/82.  
DR P-PSDB; ABP69116.  
XX  
XX  
PT New polynucleotides comprising sequences assembled from expressed  
sequence tags (ESTs), useful for treating cell-proliferative,  
neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
or coagulation disorders.  
XX  
XX  
PS Claim 1; SEQ ID NO 215; 1012pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated polynucleotide (I) comprising a  
nucleotide sequence selected from any of 948 sequences (ABZ11119-  
ABZ12066) or their mature protein coding portion, active domain coding  
protein or complementary sequences. The polynucleotides are useful for  
identifying expressed genes or for physical mapping of human genome. The  
encoded polypeptides (ABP6902-ABP6949) are useful as molecular weight  
markers, as a food supplement, for generating antibodies, in medical  
imaging, screening and diagnostic assays and for treating cell-  
proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,  
platelet or coagulation disorders, wound, burns, incision, ulcers, liver  
or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
arthritis, etc. Note: The sequence data for this patent did not form part  
of the printed specification, but was obtained in electronic format  
directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 2043 BP; 418 A; 659 C; 592 G; 374 T; 0 U; 0 Other;  
Query Match 71.6%; Score 1355.4; DB 6; Length 2043;  
Best Local Similarity 84.0%; Pred. No. 0;  
Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;  
Qy 47 CCTCCGCTGGCTCGGAGAGCGCCGCGCTGGCGAGCGGCTCATCAAGTCGCTTA 106  
Db 140 CCACCTCTGATGAGAGAGCGGCTGAAGCGGACCATCACAAACACACCTCGCGCACC 199  
Qy 107 AACCTCTGATGAGAGAGCGGCTGAAGCGGACCATCACAAACACACCTCGCGCACC 166  
Db 200 AGCCCTTATGAGAGAGCGGCTGAAGCGGACCATCACAAACACACCTCGCGCACC 259  
Qy 167 GCTACGAGTTCCTGGAGACGCTGGGCAAGGCGACCTTACGGGAGGTGAAGAGGCGACG 226  
|||||



Db 260 GCTACGAGTTCTCTGGAGACCCCTGGGCAAGGCACTTACGGGAAGGTGAAGAGCGCGGG 319  
Qy 227 AGAGCTCGGGCGCTCTGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGC 286  
Db 320 AGAGCTCGGGCGCTGTGGCCATCAAGTCAATCCGGNAGGACAAATCAAGATGAGC 379  
Qy 287 AGGATCTGCTGCACATACGGAGGAGATTGAGATCATGTCTTCACTCAACACCCCCACA 346  
Db 380 AAGATCTGATGCACATACGGAGGAGATTGAGATCATGTCTCACTCAACACCCCTACA 439  
Qy 347 TCATTGCCATCCATGAGTGTTCGAGATAGCAGCAAGATTGTGATGTCAATGAGTATG 406  
Db 440 TCATTGCCATCCATGAGTGTTCGAGAACAGCAGCAAGATCGTGTATGATGATGATG 499  
Qy 407 CCAGCCGAGGCGATCTGTATGATTACATCAGTAGCGGCCACAGGCTCAGTGAGCGGACG 466  
Db 500 CCAGCGGGCGACCTTTATGACTATCATCAGCAGCGGCAGCAGCTCAGTGAGCGGAG 559  
Qy 467 CCAGGCATTTCTCGACAGATCTGTGTCTGCCCTGCACTACTGCGACAGAACCGGATCG 526  
Db 560 CTAGGCATTTCTCGGCAGATCTGTCTGCGGTGCACTATTGCCATCAGAACAGAGTTG 619  
Qy 527 TTCAACGAGATCTCAAGCTGGAACATCTCTTAGATGCCAATGGAACATCAAGATTG 586  
Db 620 TCCACCGAGATCTCAAGCTGAGAACATCTCTTTGGATGCCAATGGAATATCAAGATTG 679  
Qy 587 CTGACTTTGGGCTCTCCAACTGTACCAAGCAAGTTCTCTCAGACGTTCTGTGGGA 646  
Db 680 CTGACTTCGGTCTCTCCAACTGTACCAATCAAGGCAAGTTCTCTGAGACATCTGTGGGA 739  
Qy 647 GCCCTCTCTACGCCCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCAGAGTTG 706  
Db 740 GCCCCCTCTATGCTCTCGCAGAGATTGTCAATGGGAAGCCCTACACAGGCCACAGGTTG 799  
Qy 707 ACAGTGTCTCTGGCGTCTCTGTATCATCTCTGTGTCATGCGACCATGCCCTTTGACG 766  
Db 800 ACAGCTGGTCTCGGTGTCTCTCTACATCTCTGGTGCATGGCACCATGCGCTTTGATG 859  
Qy 767 GGCAGGATCAAAACACTGTGTAAGCAAACTAGTAAACGGGGCTTACCGTGAGCGGCCA 826  
Db 860 GGCATGACCATAGATCTTAGTGAAACAGATCAGCAACGGGGCTTACCGGAGCCACTA 919  
Qy 827 AGCCGTTCGATGCTGTGGCTGTATCCGGTGGCTGTTAATGGTGAACCCCAACCCGTCGGG 886  
Db 920 AACCCTCTGATGCTGTGGCTGTATCCGGTGGCTGTTGATGTTGAACCCCAACCCGTCGGG 979  
Qy 887 CCACACTGGAGATGTAGCAGTCAATGTGGTGGTCAACTGGGGTTACACACCGAGTCG 946  
Db 980 CACCCCTGGAGGATGTGGCCAGTCACTGTGGTGGTCAACTGGGGCTTACGCCACCGAGTTG 1039  
Qy 947 GGGAAACAGGAAGCCCTGCTGAGGTGGGCAACCTAGTGTGACTTTGGCCGGGCTCCA 1006  
Db 1040 GAGAGAGGAGGCTCCGATAGGGGTGGGCAACCTTGGCAGTGAATCTGCCCGGCCCTCCA 1099  
Qy 1007 TGGCGGACTGTTTACGTGCTCTCTCGGCCCTCTCTGGAGAATGGAGCCAAAGTGTGCA 1066  
Db 1100 TGGCTGACTGCTCGGGCTCTCTCCGCCCTCTCTGGAGAATGGGGCCAAAGTGTGCA 1159  
Qy 1067 GCTTCTTCAAGAGCAGTGTGGGAGGTGGAAGCACTGTACCTGGGTGAGCGGCAAC 1126  
Db 1160 GCTTCTTCAAGCAGCAGTCACTGTGGTGGGGAAGCACCACCCCTGGCGTGGAGCGCCAGC 1219  
Qy 1127 ATTCTTTAAGAGTCCGGAAGGAGATGATGAGTCAAAATCTGCAAGGTGACCCCGG 1186  
Db 1220 ATTGCTTCAAGAGTCCGGAAGGAGATGATGAGTCAAAATCTGCAAGGTGACACGG 1279  
Qy 1187 CTGAGGATACCTCTTCTCGCCCTGGCAGAGCAGCTTTAAGCTTCCGAAAGGCATTCTCA 1246  
Db 1280 CTGATGACACTGCGCCATCGCCCTGGCAGAGCAACCTCAAGCTGCCAAAGGSCATTCTCA 1339  
Qy 1247 AGAAAAAGTCTCTACTCTGTCAAGGAGGAGTACAGGAGGACCTTCAGGAACTCAGACGG 1306  
Db 1340 AGAAGAGGTGTCTGAGGCTCTGAGAGAGGGGTACAGGAGGACCTTCAGGAGCTCAGCCCAA 1399

Qy 1307 TGCTGTATCTCCAGGGCAGCCTGTCTCCCTGTGTATCTCCCTGTCTCCCAAGGAAGGCATCC 1366  
Db 1400 TCCCTGCGAGCCAGGGCAGGCTGCC-----CCCTGTCTCCCAAGAGGCATTC 1450  
Qy 1367 TTAAGAGTCTCGACAGCGTGAATCTGGTTACTACTCTCTCCAGAGCCCGAGGATCTG 1426  
Db 1451 TCAAGAGCCCGACAGCGAGTCTGGCTACTACTCTCTCCGAGCCCGAGTGAATCTG 1510  
Qy 1427 GGGAACTCTTACAGCGCAGTGTGTTGTGTAGTGGGACCCCGTGGAGCAGAGTCTC 1486  
Db 1511 GGGAGCTCTTGGACGAGCGACGTGTTGTGTAGTGGGATCCAAAGAGCAGAGCCTC 1570  
Qy 1487 CACAGGCTTCAGGCGCTCTCTCCACCGCAAGGCAATCTCAAACTCAATGGCAAGTTCT 1546  
Db 1571 GCAAGCTTCAGGGCTGCTCTCTCATCGCAAGGCATCTCAAACTCAATGGCAAGTTCT 1630  
Qy 1547 CCAGCAGCCTTAGAAGGCACTACCCCTAGACACTTTGGCTCCCTGGACCAACTGGCCT 1606  
Db 1631 CCAGAGAGCCTTGGAGCTCGCGGCCCCCAACCTTCGGCTCCCTGGATGAACTGCC 1690  
Qy 1607 CTTCCCATCTGCAGCGCCGCCAGCGGCTCAGGGGCTGTGAGTGAGGACAGCATCC 1666  
Db 1691 CACTCGCCCTCTGGCCCGGCCAGCGGACCTTCAGGGCTGTGAGGAGCAGCATCC 1750  
Qy 1667 TGTCTCTCCGAGTCTTTGACCAATTTGGACTTGGCTGAAACGTCTTCCCGAAACCCCACTGA 1726  
Db 1751 TGTCTCTGAGTCTTTGACCAAGTGGACTTGGCTGAAACGTCTTCCAGAGCCCCCACTGC 1810  
Qy 1727 GGGGCTGTGTCTGTGGACAACTCAGGGGGCTTGAAGAGCCTCCCTCAAGAG----- 1780  
Db 1811 GGGGCTGTGTCTGTGGACAACTCAGGGGGCTTGAAGAGCCTCCCTCAAGAGCCCCCTG 1870  
Qy 1781 -----GTCTGAAGCGATGTTGGCGAGCAATCTTGGGGGATAGTCTCTTCTGTGACAG 1834  
Db 1871 GAAGCTGCTCTGAGCGCTTGGCGGAGATCTTGTGGGGGAGAGTCTCTTCTGTGACAG 1930  
Qy 1835 ACTGCCAAAGAGTGAAGTGAAGCTTGAAGCCTTGAAGAGCCTTAGGAATCTGTCTAAAGCTCAGC 1893  
Db 1931 ACTGCCAAGAGTGAAGTGAAGCCTTGAAGAGCCTTGAAGAGTGAAGTGAAGTGAAG 1989

RESULT 11  
ADM43851  
ID ADM43851 standard; cDNA; 2501 BP.  
XX  
AC ADM43851;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Novel human arginine-rich protein cDNA #215.  
XX  
KW ss; gene; human; arginine-rich protein; cancer; inflammation;  
KW genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN US2004053250-A1.  
XX  
PD 18-MAR-2004.  
XX  
PF 21-NOV-2002; 2002US-00302172.  
XX  
PR 05-MAR-2001; 2001US-00799451.  
PR 05-MAR-2002; 2002WO-US005095.  
PR 20-AUG-2002; 2002US-00225251.  
XX  
PA (TANG/) TANG Y T.  
PA (XUEA/) XUE A.  
PA (DRMA/) DRMANAC R T.  
XX  
Tang YT, Xue A, Drmanac RT;  
XX



DR WPI; 2004-238579/22.

XX New isolated arginine-rich protein-like polynucleotides and polypeptides, useful for diagnosing and/or treating conditions associated with aberrant PT activity of the arginine-rich polypeptides, such as cancer and inflammation.

XX

PS Disclosure; SEQ ID NO 215; 51pp; English.

XX

CC The invention relates to an isolated polynucleotide. The methods and compositions of the present invention are useful for the diagnosis and/or treatment of diseases or conditions associated with aberrant expression or activity of the arginine-rich protein-like polypeptides, such as cancer and inflammation. They can also be used in forensics, gene mapping, identification of mutations responsible for genetic disorders, and in assessing biodiversity. The present sequence represents a novel human arginine-rich protein cDNA.

XX

XX Sequence 2501 BP; 525 A; 788 C; 728 G; 460 T; 0 U; 0 Other;

XX

Query Match 71.6%; Score 1355.4; DB 12; Length 2501;

Best Local Similarity 84.0%; Pred. No. 0;

Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

Qy 47 CCTCCGCCCTCGGAGAGCGCCCGCGCTGGCGGACGGCTCATCAAGTCGCCTA 106

Db 140 CCACTCCCTCGCCGCGAGAGCTAGCCCGCGCTGGCGGAGGGCTCATCAAGTCGCCTA 199

Qy 107 AACCTCTGATGAAGAAGCAGCGGTGAAGCGGCAACCATCAAAACAACTCGCGCAC 166

Db 200 AGCCCTTAATGAAGAAGCAGCGGTGAAGCGGCAACCATCAAAACAACTCGCGCAC 259

Qy 167 GCTACGAGTTCCTCGAGCGCTGGCGGAGGCGACCTAGCGGAAGGTGAAGAAGCAG 226

Db 260 GCTACGAGTTCCTCGAGCGCTGGCGGAGGCGACCTAGCGGAAGGTGAAGAAGCG 319

Qy 227 AGAGCTCGGGCGCTGTGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGC 286

Db 320 AGAGCTCGGGCGCTGTGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGC 379

Qy 287 AGAGTCTGTGCATACGAGGAGAGTTGAGATCATGTCTTCACTCAACACCCCA 346

Db 380 AAGATCTGTGCATACGAGGAGAGTTGAGATCATGTCTTCACTCAACACCCCA 439

Qy 347 TCATTGCCATCCATGAAGTGTTCAGAAATACGACAGAGATGTGATGTATGAGTATG 406

Db 440 TCATTGCCATCCATGAAGTGTTCAGAAATACGACAGAGATGTGATGTATGAGTATG 499

Qy 407 CCAGCCGAGGCGATCTGTATGATATACATCATGAGCGGCCACCGCTCAGTGAGCGGACG 466

Db 500 CCAGCCGAGGCGATCTTATGATATACATCATGAGCGGCCACCGCTCAGTGAGCGGACG 559

Qy 467 CCAGGCAATTTCTCCGACAGATGTGTGCGCTGCACTACTGCCACAGAACGGGATCG 526

Db 560 CTAGGCAATTTCTCCGACAGATGTGTGCGCTGCACTACTGCCACAGAACAGATTG 619

Qy 527 TTACCCGAGATCTCAAGCTGGAACATCTTCTAGATGCCAATGGAACATCAAGATTG 586

Db 620 TCACCCGAGATCTCAAGCTGGAACATCTTCTAGATGCCAATGGAACATCAAGATTG 679

Qy 587 CTGACTTTGGCCTCTCAACCTGTACCAAGGCAAGTTCTCTCAGACGTTCTGTGGGA 646

Db 680 CTGACTTTGGCCTCTCAACCTGTACCAAGGCAAGTTCTCTCAGACATCTGTGGGA 739

Qy 647 GCCCTCTCTAGCCTCGCTGAGATGATGACAGGGAAGCCCTATGTGGGCCAGAGTGG 706

Db 740 GCCCTCTCTAGCCTCGCTGAGATGATGACAGGGAAGCCCTATGAGGAGGAGTGG 799

Qy 707 ACAGCTGCTCTGTGGGCGTCTCTCTGTACATCTCTGTGTGATGACCATCCCTTTGACG 766

Db 800 ACAGCTGCTCTGTGGGCGTCTCTCTGTACATCTCTGTGTGATGACCATCCCTTTGATG 859

Qy 767 GGCAGGATCAATAAACAACCTGTGAAGCAATCAAGTAAACGGGGCTTACCGTGAGCGGCCA 826

Db 860 GGCATGACCAATAGATCTTAGTGAACAGATCAGAACGGGGCTTACCGGAGCCACTA 919

Qy 827 AGCCGTCCGATGCTGTGGCCTGATCCGCTGGCTGTAAATGTGTGAACCCACCCGTCGG 886

Db 920 AACCTCTGATGCTGTGGCCTGATCCGCTGGCTGTAAATGTGTGAACCCACCCGTCGG 979

Qy 887 CACACTGGAGGATGATAGCCAGTCAATTTGGGTCAACTGGGGTTACACACCGAGTCG 946

Db 980 CCACTCTGGAGGATGATGGCCAGTCACTGGTGGGTCACTGGGGCTAGCCACCCGAGTGG 1039

Qy 947 GCGAACAGGAAGCCCTCGCTGAGGGTGGGCAACCTAGTGTGACTTTGGCCGGGCTCCCA 1006

Db 1040 GAGAGCAGGAGGCTCCGCATGAGGGTGGGCAACCTGTGAGTGTCTGCCCCGGCTCCCA 1099

Qy 1007 TGCGGAGATGTTTACGTGCTCTCGCGCCCTCTCTGGAGAAATGGAGCAAGGTGTGCA 1066

Db 1100 TGGCTGACTGGCTCCGGCGTTCTCCCGCCCTCTCTGGAGAAATGGGCAAGGTGTGCA 1159

Qy 1067 GCTTCTTCAAGCAGCAGTCCCGGAGGTGAAGCACTGTACTGGCTGGAGCGGCAAC 1126

Db 1160 GCTTCTTCAAGCAGCAGTCCCGGAGGTGAAGCACTGTACTGGCTGGAGCGGCAAC 1219

Qy 1127 ATTCTCTTAAGAAAGTCCCGAAAGAGATGACATGGCTCAAAATCTGCAAGGTGACCCGG 1186

Db 1220 ATTCTCTTAAGAAAGTCCCGAAAGAGATGACATGGCTCAAAATCTGCAAGGTGACCCGG 1279

Qy 1187 CTGAGGATACCTTTCTCGCCCTGGCAAGAGCAGCTTAAAGCTTCGAAAGGATTTCTCA 1246

Db 1280 CTGATGACATGCCCCATCGCCCTGGCAAGAGCAACCTCAAGCTGCAAGGGCATTTCTCA 1339

Qy 1247 AGAAAGTCTCTACCTCGTCAGGGAGGTACAGGAGACCTCAGGAACCTCAGACCGG 1306

Db 1340 AGAAAGTCTCTACCTCGTCAGGGAGGTACAGGAGACCTCAGGAACCTCAGGCGCA 1399

Qy 1307 TGCTCTGATCTCCAGGCGACCTGCTCCCTGTATCTCTCTCTCCCAAGAAAGGATCTC 1366

Db 1400 TCCTCTGAGCCAGGCGAGGCTGCCCC-----CCCTGCTCCCAAGAGGATCTC 1450

Qy 1367 TTAAGAAAGTCTCGACAGCGTGAATCTGGTTACTTCTCTCCAGAGCCAGCGATCTG 1426

Db 1451 TCAAGAAAGCCCGACAGCGAGTCTGGCTACTTCTCTCCAGAGCCAGTGAATCTG 1510

Qy 1427 GGAAGTCTTAGAGCGCAGTGTGTTGTGAGTGGGATCCCAAGGAGCAGAACTCTC 1486

Db 1511 GGAAGTCTTAGAGCGCAGGCGAGTGTGTTGTGAGTGGGATCCCAAGGAGCAGAACTCTC 1570

Qy 1487 CACAGGCTTCAGGCTCTCTCTCCACCGCAAGGGCAATTCCTCAAACTCAATGGCAAGTTCT 1546

Db 1571 CGCAAGCTTCAGGCTCTCTCTCCATCCGCAAGGGCAATTCCTCAAACTCAATGGCAAGTTCT 1630

Qy 1547 CCGGCAAGCTTTAGAGGCACTACCCCTAGACACTTTGGCTCTCTCGAGCCAGCGCTC 1606

Db 1631 CCAGAGCAGCTTTGAGCTCGCGGCCCCCAACCACTTCGCTCTCTCGAGTGAATCTGCCCC 1690

Qy 1607 CTTCCCATCTCGAGCCCGCCCGCCAGCGCTCAGGGCTGTGAGTGGAGCAGCATCC 1666

Db 1691 CACTCGCCCCCTGGCCCGGCGGCGAGCCCTCAGGGGCTGTGAGGAGCAGCATCC 1750

Qy 1667 TGTCTCTCCAGTCTTTTGAACCAATTTGGACTTGTGCTGAAGCTCTTCCCGCAACCCCACTGA 1726

Db 1751 TGTCTCTCTGAGTCTTTTGAACAGTGTGCTTGTGCTGAACGGCTCCAGAGCCCGCACTGC 1810

Qy 1727 GGGGCTGTGTGTGTGGAACAACCTGAGGGGCTTGAAGAGCCCTCCCTCAGAG----- 1780

Db 1811 GGGGCTGTGTGTGTGGAACAACCTCAGGGGCTTGAAGAGCCCGCTCAGAGGGCCCTG 1870

Qy 1781 -----GTCTGAAGCGATGTTGGCAGAAATCTTTGGGGATAGTCTTTTCTCAGCAG 1834

Db 1871 GAAGCTGCTTGAAGCGCTGTGGCAGGATCTTTTGGGGAGCAGCTGCTTTTCCCTGACAG 1930

Qy 1835 ACTGCCAAGAGGTGATGTCAGGCTTACAGAAAGCCCTAGGAATCTGCTCAAGCTCAGC 1893

Db 1931 ACTGCCAGGAGGTGACAGCGACCTACCGACAGGCACTGAGGGTCTGTCTCAAAGCTCACC 1989

RESULT 12  
ABX71420  
ID ABX71420 standard; cDNA; 3443 BP.  
XX AC ABX71420;  
XX DT 14-APR-2003 (first entry)  
XX DE Human cell cycle-associated cDNA from clone DKFZphtes3\_7j3.  
XX KW Human; gene; gene therapy; vaccine; disease treatment; detection; ss.  
XX OS Homo sapiens.  
XX PN WO200112659-A2.  
XX PD 22-FEB-2001.  
XX PF 18-AUG-2000; 2000WO-IB001496.  
XX PR 18-AUG-1999; 99US-0149499P.  
XX PR 28-SEP-1999; 99US-0156503P.  
XX PA (GHEU-) GERMAN HUMAN GENOME PROJECT.  
XX PI Wiemann S;  
XX DR WPI; 2001-327840/34.  
XX DR P-PSDB; ABUS3319.  
XX PT Nucleic acids having the sequences of clones isolated from libraries of  
XX PT different human tissues, useful in recombinant DNA methodologies.  
XX PS Claim 1; Page 942-943; 1095pp; English.  
XX CC This invention describes novel polynucleotides and polypeptides isolated  
XX CC from human cDNA libraries which can be used for gene therapy or in  
XX CC vaccines. The polynucleotides of the invention and antibodies encoded by  
XX CC them may be used in the prevention, diagnosis and treatment of diseases  
XX CC associated with inappropriate polypeptide expression. The products of the  
XX CC invention may also be used to identify modulators of expression and  
XX CC activity and to down regulate expression and activity. The antibodies of  
XX CC the invention may also be used as diagnostic agents for detecting the  
XX CC presence of polypeptides in samples. This sequence encodes a polypeptide  
XX CC described in the disclosure of the invention  
SQ Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;

Query Match 71.6%; Score 1355.4; DB 5; Length 3443;  
Best Local Similarity 84.0%; Pred. No. 0;  
Matches 1562; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

Qy 47 CCTCGCGCTGGCTCGAGAGCGCCCGCGCTGGCGGACGGGCTCATCAAGTCGCCTA 106  
Db 162 CCATCTCCTCGCGGAGAGCTAGCCCGCGCTGGCGGAGGGCTGATCAAGTCGCCCA 221  
Qy 107 AACCTCTGATGAAGAAGCAGCGGTGAAGCGGCACCATCAAAACACAACCTGCGGCACC 166  
Db 222 AGCCCTTAATGAAGAAGCAGCGGTGAAGCGGCACCATCAAAACACAACCTGCGGCACC 281  
Qy 167 GCTACGAGTTCCTGGAGACCTGGGCAAGGCACTAGGGAGGTGAAGAGCAGAG 226  
Db 282 GCTACGAGTTCCTGGAGACCTGGGCAAGGCACTAGGGAGGTGAAGAGGCGCGGG 341  
Qy 227 AGAGCTCGGGCGCTCTGGTGGCCATCAAGTCCATCAGAAAGACAAAATCAAGATGAGC 286  
Db 342 AGAGCTCGGGCGCTCTGGTGGCCATCAAGTCCATCAGAAAGACAAAATCAAGATGAGC 401  
Qy 287 AGGATCTGCTGCACATACGGAGGAGATTGAGATCATGTCTTCACTCAACACCCCCACA 346

Db 402 AAGATCTGATGSCATACGAGGGAGATTGAGATCATGTCTCATCACTCAACACCCCTCACA 461

Qy 347 TCATTGCCATCCATGAAGTCTTTGAGAAATAGCAGCAAGATTGTGATTGTCTATGAGTATG 406

Db 462 TCATTGCCATCCATGAAGTCTTTGAGAAATAGCAGCAAGATTGTGATTGTCTATGAGTATG 521

Qy 407 CCAGCCGAGCGGATCTGTATGATTATCATAGTGAAGCGCCACCGCTGAGTGAAGCGGACG 466

Db 522 CCAGCCGAGCGGATCTGTATGATTATCATAGTGAAGCGCCACCGCTGAGTGAAGCGGACG 581

Qy 467 CCAGCGATTTCTTCGACAGATCGTGTCTGCGCTGCACTACTGCCACAGAAACGGGATCG 526

Db 582 CTAGGCAATTTCTTCGCGCAGATCGTCTCTGCGTGCATATTGCTCATCAGAACAGAGTTG 641

Qy 527 TTCAACGAGATCTCAAGCTGGAACAACTCTCTAGATGCCAATGGAACAACTCAAGATTG 586

Db 642 TCACCCGAGATCTCAAGCTGGAACAACTCTCTAGATGCCAATGGAACAACTCAAGATTG 701

Qy 587 CTGACTTTGGCTCTCTCAACCTGTACCAAGGCAAGTTCTCTCCAGACGTTCTGTGGGA 646

Db 702 CTGACTTTGGCTCTCTCAACCTGTACCAAGGCAAGTTCTCTCCAGACGTTCTGTGGGA 761

Qy 647 GCGCTCTCTAGGCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCAGAGGTGG 706

Db 762 GCGCCCTCTATGCTTCGCCAGAGATTGTCAATGGGAAGCCCTACACAGAGCCACAGAGTGG 821

Qy 707 ACAGCTGCTCTCTGGGGTCTCTCTGTACATCTCTGTGTGATGACCACTGCGCTTTGACG 766

Db 822 ACAGCTGCTCTCTGGGGTCTCTCTGTACATCTCTGTGTGATGACCACTGCGCTTTGATG 881

Qy 767 GGCAGGATCATAAACACTGCTGAAAGCAATCAAGTAAACGGGGCTTACCGTGAAGCGGCCA 826

Db 882 GGCATGACCAATAGATCTTAGTGAACAGATCAGCAACGGGGCTTACCGGAGGCCACTA 941

Qy 827 AGCGCTCCGATGCTGTGGCTGATCGGTGGCTGTAAATGTTGTAACCCACCGCTCGGG 886

Db 942 AACCTCTGATGCTCTGGGCTGATCGGTGGCTGTAAATGTTGTAACCCACCGCTCGGG 1001

Qy 887 CCACACTGAGGAGTGTAGCCAGTCAATTGGTGGGTCAACTGGGGTTACACACCGAGTCG 946

Db 1002 CCACCTTGAAGGATGTGGCCAGTCACTGTGGGTCAACTGGGGTTAGCCACCGAGTGG 1061

Qy 947 GGAACAGGAAGCCCTCGCTGAGGGTGGGCAACCTAGTGGTGAATTTGGCGGGCTCTCA 1006

Db 1062 GAGAGCAGGAGGCTCCGCTGAGGGTGGGCAACCTGSCAGTACTCTGCCCGGCTCTCA 1121

Qy 1007 TGGCGACTGTTAGTGTGCTCTCTCGGCCCTCTCTCGAGATGAGAACCAAGGTGTGCA 1066

Db 1122 TGGCTGACTGGCTCCGCGCTTCTCTCGGCCCTCTCTCGAGAAATGGGGGCCAAGGTGTGCA 1181

Qy 1067 GCTTCTTCAAGCAGCAGTGTCCGGAGGTGGAAGCACTGTACCTGGGCTGGAGCGGCAAC 1126

Db 1182 GCTTCTTCAAGCAGCAGTGTCCGAGGTGGGGAAGCACTACCTCTGGCTGGAGCGGAGC 1241

Qy 1127 ATTCTCTTAAGAAAGTCCGAAAGAGAAATGACATGGCTCAAAATCTCAAGAGTCAACCGG 1186

Db 1242 ATTGCTCAAGAGTCCGCAAGAGAAATGACATGGGCCAGTCTCTCCACAGTACACGG 1301

Qy 1187 CTGAGGATACCTCTTCTCGCCCTGGCAAGAGAGCCCTTAAGCTTCGAAAGGCAATTTCA 1246

Db 1302 CTGATGACATGCCCCCTCGCCCTGGCAAGAGCAACCTCAAGCTGCAAGAGGCAATTTCA 1361

Qy 1247 AGAAAGTCTCTACTCTGTGAGGGGAGGTACAGGAGGACCTCAGGAACTCAGACCGG 1306

Db 1362 AGAAGAGGTTGTGAGGCTCTGCAGAAAGGGGTACAGGAGGACCTCTCGGAGCTCAGCCAA 1421

Qy 1307 TGCTCTGATCTCCAGGCGAGCTGTCTCTGTCTGTATCTCTCTCTCTCTCTCTCTCTCTCT 1366

Db 1422 TCCCTGCGAGCCCGAGGCGAGCTGCCC-----CGCTGCTCTCTCTCTCTCTCTCTCTCT 1472

Qy 1367 TTAAGAGTCTCGACAGCGTGAATCTGTTTACTACTCTCTCTCTCTCTCTCTCTCTCTCTCT 1426

Db 1473 TCAAGAGCCCGCAGCAGCGAGTCTGGCTACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1532





Db 282 GCTACGAGTTCTCTGGAGACCTTGGGCAAGGACCTTACGGGAAGGTGAAGAAGCGCGGG 341  
Qy 227 AGAGCTCGGGCGTCTGGTGGCCATCAAGTCCATCAGGAAGAAGCAAAATCAAGATGAGC 286  
Db 342 AGAGCTCGGGCGCTGGTGGCCATCAAGTCAATTCGGGAAGGACAAATCAAGATGAGC 401  
Qy 287 AGAGTCTGCTGCACATACGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCA 346  
Db 402 AAGATCTGATGCACATACGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCA 461  
Qy 347 TCATTGCCATTCATGAAGTGTTCAGAAATAGCAGCAAGATTGTGATTGTATGATGATG 406  
Db 462 TCATTGCCATTCATGAAGTGTTCAGAAATAGCAGCAAGATTGTGATTGTATGATGATG 521  
Qy 407 CCAGCCGAGGGATCTGTATGATTACATCAGTAGAGCGGCCACCGCTGAGTGAGCGGAGC 466  
Db 522 CCAGCCGAGGGCGACCTTTATGACTACATCAGCGAGCGCGCAGCAGCTCAGTGAGCGCAAG 581  
Qy 467 CCAGGCAATTTCTCCGACAGATCGTGTCTGCCCTGCACTACTGCCACACAGAAAGGATCG 526  
Db 582 CTAGGCAATTTCTCCGCGAGATCGTCTGCGCGTGCACTATTGCCATCAGAAACAGATTG 641  
Qy 527 TTCAACGAGATCTCAAGCTGGAACATCTCTTCTAGATGCCAATGGAACATCAAGATTG 586  
Db 642 TCCACCGAGATCTCAAGCTGGAACATCTCTTCTGATGCCAATGGAATATCAAGATTG 701  
Qy 587 CTGACTTTGGCTCTCAACCTGTACCAAGCAAGTTCTCTCCAGACGTTCTGTGGGA 646  
Db 702 CTGACTTTGGCTCTCAACCTGTACCAAGCAAGTTCTCTCCAGACATTTCTGTGGGA 761  
Qy 647 GCCCTCTCTAGCGCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGCCCGCAGAGTGG 706  
Db 762 GCCCTCTCTAGCGCTCGCTGAGATAGTCAATGGGAAGCCCTACAGAGCCAGAGTGG 821  
Qy 707 ACAGCTGGTCTCTGGGCGTTCTCCTGTACATCTCTGGTGCATGCGACCATGCTTTGACG 766  
Db 822 ACAGCTGGTCTCTGGGCGTTCTCCTGTACATCTCTGGTGCATGCGACCATGCTTTGATG 881  
Qy 767 GGCAGGATCATAAACACTGGTGAAGCAATCAGTAAAGGGGCTTACGGTAGCGCGCCCA 826  
Db 882 GGCATGACCATAGATCCTAGTGAACAGATCAGCAACGGGGCCCTACCGGAGGCCACTA 941  
Qy 827 AGCGCTCCGATGCTGTGGCTGATCGGTGGCTGTTAAATGGTGAACCCCAACCGCTCGGG 886  
Db 942 AACCTCTGATGCTGTGGCTGATCGGTGGCTGTTGATGGTGAACCCCAACCGCGCGG 1001  
Qy 887 CCACACTGGAGGATGTAGCCAGTCAATGGTGGGTCAA CTGGGTTTACCAACCGAGTGC 946  
Db 1002 CCACCTGGAGGATGTGGCCAGTCACTGGTGGGTCAA CTGGGCTTACGCCACCGAGTGG 1061  
Qy 947 GGGNAACAGGAGCCCTGGGTGAGGGTGGGACCCCTAGTGTGACTTTGGCGGGCTTCCA 1006  
Db 1062 GAGACGAGGAGGCTCCGATGAGGGTGGGACCCCTGGCAGTGAATCTGCGCGGCTTCCA 1121  
Qy 1007 TGGCGGACTGGTACGTCTGCTCGCGCCCTCTCTGGAGAAATGGAGCAAGGTGTGCA 1066  
Db 1122 TGGCTGACTGGCTCGGGGTCTCTCCGCCCTCTCTGGAGAAATGGGCAAGGTGTGCA 1181  
Qy 1067 GCTTCTTCAAGCAGCAGCTGCGGGAGGTGGAAGCACTGTACTGGGTGAGAGCGGCAAC 1126  
Db 1182 GCTTCTTCAAGCAGCAGCTGCGGTGGGGAAGCAACACCCCTGGCTGGAGCGCCAGC 1241  
Qy 1127 ATTCTCTTAAGAGTCCGGAAGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGG 1186  
Db 1242 ATTCTCTTAAGAGTCCGGAAGAGAAATGACATGGCTCAAAATCTGCAAGGTGACACGG 1301  
Qy 1187 CTGAGGATACCTCTTCTCGCCCTGGCAAGAGCAGCTTAAAGTTCGGAAGGCAATCTCA 1246  
Db 1302 CTGATGACATGCCCCCTGGCAAGAGCAATCTCAAGTCTGCAAGGCAATCTCA 1361  
Qy 1247 AGAAAAAGTCTCTCTCTGTCAGGGAGGTGTACAGGAGGACCTCTCAGAACTCAGACCGG 1306  
Db 1362 AGAAGAGGTGTACGCTCTCTGCAAGAGGGGTACAGGAGGACCTCTCGGAGCTCAGCCAA 1421

Qy 1307 TGCCTGATATCTCAGGGCAGCCTGTCCCTGCTGTATCCCTGCTCCAGGAAGGCAATCC 1366  
Db 1422 TCCCTTGGAGCCAGGGCAGGCTGCCC-----CGTGTCTCCCAAGAAAGGCAATC 1472  
Qy 1367 TTAAGAAAGTCTCGACAGCGTGAATCTGGTTACTACTCTCTCCAGAGCCCGAGGCTGTG 1426  
Db 1473 TCAAGAAAGCCCGACAGCGGAGTCTGGCTACTCTCTCTCCGAGCCCGAGTGAATCTG 1532  
Qy 1427 GGAACCTTTAGACCGCAGTGATGTGTGTGTAGTGGGACCCCGTGGAGCAGAACTCTC 1486  
Db 1533 GGGAGCTCTTGGACGCGAGCGAGTGTGTGTGTAGTGGGATCCCAAGGAGCAGAAAGCTC 1592  
Qy 1487 CACAGGCTTCAGGGCTCTCTCCACCGCAGGAGGCAATCTCAAACTCAATGGCAAGTTCT 1546  
Db 1593 CGCAAGCTTCAGGGCTGCTCTCCATCGCAAGGCAATCTCAAACTCAATGGCAAGTTCT 1652  
Qy 1547 CCGCACAGCCTTTAGAAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCCT 1606  
Db 1653 CCAGACAGCCTTTGGAGCTCGCGGCCCCACACCTTGGCTCCCTGGATGAATCGCCC 1712  
Qy 1607 CTTCCCATCTGACGCGCGGCCCGCCCTCAGGGGCTGTGAGTGAGGACAGCAATCC 1666  
Db 1713 CACTCGCCCCCTGGCCCGGGCCAGCGACCTCAGGGGCTGTGAGCGAGGACAGCATCC 1772  
Qy 1667 TGTCTCCGAGTCTTTGACCAATTTGCACTTGCCTGAACTCTTCCGAAACCCCACTGA 1726  
Db 1773 TGTCTCTGAGTCTTTGACCACTTGGAGTGTGCTGAACGGCTCCAGAGCCCCCACTGC 1832  
Qy 1727 GGGGCTGTGTGTCTGTGGCAACCTCAGGGGGCTTGAAGCAGCTCCCTCAGAAAG----- 1780  
Db 1833 GGGGCTGTGTGTCTGTGGCAACCTCAGGGGCTTGAAGGAGCCCCCTCAGAGGGCCCTG 1892  
Qy 1781 -----GTCTGAACGAGTGTGCGAAGTCTTGGGGGATAGTGTCTTTCTGTGACAG 1834  
Db 1893 GAAGCTGCTTGAAGCGCTGGCGGCGAGATCCTTTTGGGGGACAGCTGCTTTTCCCTGACAG 1952  
Qy 1835 ACTGCCAAGAGGTGACTGCGAGCTCAGACAAAGCCCTAGGAATCTGCTCAAGCTCAGC 1893  
Db 1953 ACTGCCAAGAGGTGACAGCGACCTTACGACAGGCACTGAGGGTCTGCTCAAGCTCACC 2011

## RESULT 15

ADO20171  
ID ADO20171 standard; cDNA; 3443 BP.

XX AC ADO20171;

XX AC ADO20171;

DT 12-AUG-2004 (first entry)

XX Human PRO polynucleotide #540.

DE Human; PRO; gene; ss; immune related disorder;

XX Human; PRO; gene; ss; immune related disorder;

KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;

KW juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;

KW vasculitis; sarcoidosis; autoimmune haemolytic anaemia;

KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;

KW renal disease; demyelinating disease; central nervous system;

KW peripheral nervous system; demyelinating polyneuropathy;

KW Guillain-Barre syndrome;

KW chronic inflammatory demyelinating polyneuropathy.

XX Homo sapiens.

OS Homo sapiens.

XX WO200403361-A2.

XX 27-MAY-2004.

XX 06-NOV-2003; 2003WO-US035268.

XX 08-NOV-2002; 2002US-0425235P.

XX (GETH ) GENENTECH INC.

XX (GETH ) GENENTECH INC.

XX (GETH ) GENENTECH INC.

XX (GETH ) GENENTECH INC.

XX (GETH ) GENENTECH INC.

XX (GETH ) GENENTECH INC.

XX (GETH ) GENENTECH INC.

XX (GETH ) GENENTECH INC.

XX (GETH ) GENENTECH INC.

XX (GETH ) GENENTECH INC.

XX (GETH ) GENENTECH INC.

XX (GETH ) GENENTECH INC.

XX	Fong S,	Dennis K,	Clark H,	Chiu H,	Schoenfeld J,	Williams PM;
PI	Wood WI,	Wu TD;				
XX	WPI; 2004-420067/39.					
DR	P-PSDB; ADO20172.					
DR						
XX	Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for					
PT	treating an immune related disorder such as systemic lupus erythematosus,					
PT	rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or					
PT	spondyloarthritis.					
XX						
PS	Claim 1; SEQ ID NO 1136; 1731pp; English.					
XX						
CC	The invention relates to human PRO polypeptides and the polynucleotides					
CC	encoding them. The polypeptides and polynucleotides are useful for					
CC	treating and diagnosing immune related disorders in mammals. The immune					
CC	related disorders include systemic lupus erythematosus, rheumatoid					
CC	arthritis, osteoarthritis, juvenile chronic arthritis, systemic					
CC	sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune					
CC	haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes					
CC	mellitus, immune-mediated renal disease, demyelinating diseases of the					
CC	central or peripheral nervous system, demyelinating polyneuropathy,					
CC	Guillain-Barre syndrome and chronic inflammatory demyelinating					
CC	polyneuropathy. This sequence represents a human PRO polynucleotide of					
CC	the invention.					
XX						
SEQ	Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;					
	Query Match	71.6%;	Score 1355.4;	DB 12;	Length 3443;	
	Best Local Similarity	84.0%;	Pred. No. 0;			
	Matches 1562;	Conservative	0;	Mismatches 276;	Indels 21;	Gaps 2;
QY	47	CCTCGGCCCTGCGCTCGAGAGCGCCGGCGCTGGCGACGGGTGCATCAAGTCGCGTA	106			
DB	162	CCATCTCCCTGCGCCGACAGCTAGCCCGGCCCTGGCGAAGGGTGATCAAGTCGCCCA	221			
QY	107	AACCTTCTGATGAAGAAGCAGCGCGTGAAGCGGCACCATCAAAACAACACTTCGCGCAC	166			
DB	222	AGCCCCAATGAAGAAGCAGCGCGTGAAGCGSCACCACCACAAGCACACCTTCGCGCAC	281			
QY	167	GCTAGAGTTCTTGAGACGCTGGCAAGGGCACCTACGGGAAGGTGAAGAAGGCACGAG	226			
DB	282	GCTACGAGTTCTTGAGACCTTGGGCAAGGCACCTACGGGAAGGTGAAGAAGGGCGGG	341			
QY	227	AGAGTCGGGGCGTCTGTGGCCATCAAGTCCATCAGGAAGACAAAATCAAGATGAGC	286			
DB	342	AGAGTCGGGGCGCTGTGGCCATCAAGTCAATCCGAAGGACAAAATCAAGATGAGC	401			
QY	287	AGGATCTGCTGCACATACGGAGGGAGATTGAGATCATGTCTTCACTCAACCAACCCCACA	346			
DB	402	AAGATCTGATGCACATACGGAGGGAGATTGAGATCATGTTCATCTCAACCAACCTTCCA	461			
QY	347	TCAITGGCCATCCATAAGTGTGTGAAATAGACAGCAAGATTGTGATTGTCTGAGGATG	406			
DB	462	TCATTGCCATCCATCAAGTGTGTGAAACAGCAGCAAGATCGTGATCGTCTATGGAGTAG	521			
QY	407	CCAGCCGAGGGCATCTGATATCATCATCATGAGCGGCCACCGCTGAGTGAGCGGGACG	466			
DB	522	CCAGCCGAGGGGACCTTTATGATACATCATCAGCAGCGGCAGCAGCTCATGTGAGCGCGAAG	581			
QY	467	CCAGGCATTTCTTCGACAGATCGTGTCTGCCCCTGCATCTACTGCCACACAGAACGGATCG	526			
DB	582	CTAGGCATTTCTTCGCGCAGATCGTCTCTGCGGTGCATCTTTGCCATCNGAACAGATTG	641			
QY	527	TTACCGAGATCTCAAAGCTGGAAAAATCATCTTTCTAGATGCCAATGGAAACATCAAGATTG	596			
DB	642	TCCACCGAGATCTCAAGCTGAGAAACATCCTCTTGGATGCCAATGGGAATCAAGATTG	701			
QY	587	CTGACTTTGGCTCTCCAACTGTACCAAGGCAAGTTCTCTCAGACGTTCTCTGTGGGA	646			
DB	702	CTGACTTTGGCTCTCCAACTGTACCAAGGCAAGTTCTCTGACAGACATTTCTGTGGGA	761			

Db	1833	GGGGCTGTGTCTGTGTGGACAACCTCAGGGGCTTGAGGAGCCCCCTCAGAGGCCCTG	1892
Qy	1781	-----GTCTGAAGCGATGTGGCAGGAATCCTTTGGGGGATAGCTGCTTTTCTGTGACAG	1834
Db	1893	GAAAGCTGCCTGAGGGCGCTGGCGGAGGATCCTTTGGGGGACAGCTGCTTTTCCCTGACAG	1952
Qy	1835	ACTGCCAAGAGGTGACTGCGAGCCTACAGACAAGCCCTAGGAATCTGCTCAAAGCTCAGC	1893
Db	1953	ACTGCCAGGAGGTGACAGGACCTTACCGACAGGCACTGAGGGTCTGCTCAAAGCTCACC	2011

Search completed: January 24, 2005, 16:12:26  
 Job time : 883.636 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 14:50:33 ; Search time 7944.68 Seconds  
(without alignments)  
11267.845 Million cell updates/sec

Title: US-09-980-464-4\_COPY\_123\_2015

Perfect score: 1893  
Sequence: 1 atggagtcggtgccttact.....gaatctgctcaagtcagc 1893

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1893	100.0	2902	6	AR279568 Sequence
2	1888.2	99.7	2027	6	AX381046 Sequence
3	1872.2	98.9	2917	10	BC033302 Mus muscu
4	1859	98.2	2938	10	BC046833 Mus muscu
5	1683.4	88.9	2026	6	AX380986 Sequence
6	1683.4	88.9	2929	6	AX380960 Sequence
7	1528.4	80.7	291762	2	AC125887 Rattus no
8	1360.2	71.9	1884	6	AX407118 Sequence
9	1360.2	71.9	3353	6	AX407116 Sequence
10	1357.6	71.7	3360	6	AX642961 Sequence
11	1357	71.7	1887	6	CQ714327 Sequence
12	1357	71.7	3395	6	CQ782778 Sequence
13	1357	71.7	3395	6	BD127309 Primer fo
14	1357	71.7	3395	9	AK074830 Homo sapi
15	1357	71.7	3402	9	BC017306 Homo sapi
16	1355.4	71.6	3443	6	AX086949 Sequence
17	1355.4	71.6	3443	6	AX780351 Sequence
18	1355.4	71.6	3443	9	HSM801859 Homo sapi
19	1338.8	70.7	2291	6	AX056394 Sequence

20	1230	65.0	1833	6	AX803413 Sequence
21	1071.4	56.6	1186	6	AX380972 Sequence
22	943.4	49.8	1186	6	AX381029 Sequence
23	923.4	48.8	252978	2	AC126983 Rattus no
24	840.2	44.4	2616	6	AX430369 Sequence
25	840.2	44.4	2619	6	AX430468 Sequence
26	723	38.2	179414	9	AC098933 Homo sapi
27	719.8	38.0	6389	9	HSM806181
28	651.8	34.4	1274	5	BX950299 Gallus ga
29	648	34.2	707	6	AX381050 Sequence
30	648	34.2	712	6	AX380990 Sequence
31	644.2	34.0	661	6	AX381052 Sequence
32	613.6	32.4	741	6	AX380991 Sequence
33	592	31.3	734	6	AX381051 Sequence
34	583.2	30.8	588	6	AX381056 Sequence
35	557.4	29.4	660	6	AX380992 Sequence
36	533	28.2	578	6	AX381054 Sequence
37	515	27.2	548	6	AX381055 Sequence
38	510.6	27.0	579	6	AX380994 Sequence
39	506.4	26.8	1986	6	AX213158 Sequence
40	506.4	26.8	2884	6	AX213160 Sequence
41	506.4	26.8	6825	6	CQ715220 Sequence
42	506.4	26.8	6828	6	CQ776374 Sequence
43	506.4	26.8	6828	6	AX213156 Sequence
44	506.4	26.8	6828	9	AB011109 Homo sapi
45	493.6	26.1	585	6	AX380996 Sequence

ALIGNMENTS

RESULT 1  
AR279568  
LOCUS AR279568 2902 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 4 from patent US 6514719.  
ACCESSION AR279568  
VERSION AR279568.1 GI:29714427  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2902)  
AUTHORS Bird,T.A., Virca,G.D., Martin,U. and Anderson,D.M.  
TITLE Methods for identifying compounds that alter kinase activity  
JOURNAL Patent: US 6514719-A 4 04-FEB-2003;  
FEATURES  
source Location/Qualifiers  
1..2902  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match	100.0%;	Score 1893;	DB 6;	Length 2902;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1893;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGGAGTCGGTGGCTTACTCCAGCGCCGAGCGAGCTCCCTCGGCTCGCGCTCGGCC	60	
Db	123	ATGGAGTCGGTGGCTTACTCCAGCGCCGAGCGAGCTCCCTCGGCTCGCGCTCGGCC	182	
Qy	61	TCGGAGAGCGCCCGCGCTGGCGGACGGGCTCATCAAGTCGCTAAACCTCTGATGAAG	120	
Db	183	TCGGAGAGCGCCCGCGCTGGCGGACGGGCTCATCAAGTCGCTAAACCTCTGATGAAG	242	
Qy	121	AAGCAGCGGTGAAGCGGACCATCAACAACAACTCGGCGACCGCTACGAGTTCCTG	180	
Db	243	AAGCAGCGGTGAAGCGGACCATCAACAACAACTCGGCGACCGCTACGAGTTCCTG	302	
Qy	181	GAGACCTGGGCAAGGCGACCTACGGAAGGTGAAGAAGCAGCAGAGAGCTCGGGCGCT	240	
Db	303	GAGACCTGGGCAAGGCGACCTACGGAAGGTGAAGAAGCAGCAGAGAGCTCGGGCGCT	362	
Qy	241	CTGGTGGCCATCAAGTCCATCAGGAAGACAAAATCAAGATGAGCAGAGTCTGCTGCAC	300	

Db 363 CTGTGGCCATCAAGTCCATCAGGAAGACAAAATCAAGATGACAGGATCTGCTGCAC 422  
Qy 301 ATACGGAGGAGATTGAGATCATGTCTTCACTCAACACACCCCCACATCAITTTGCCATCAT 360  
Db 423 ATACGGAGGAGATTGAGATCATGTCTTCACTCAACACACCCCCACATCAITTTGCCATCAT 482  
Qy 361 GAACTGTTTGAATATAGCAGCAAGATTCTGATTTGTTCATGGAGTAGTCCAGCCGAGGCGAT 420  
Db 483 GAACTGTTTGAATATAGCAGCAAGATTCTGATTTGTTCATGGAGTAGTCCAGCCGAGGCGAT 542  
Qy 421 CTGTATGATTATCATCAGTAGCGGCCACCGCTGAGTGAAGCGGACCGCCAGGCAATTTCTTC 480  
Db 543 CTGTATGATTATCATCAGTAGCGGCCACCGCTGAGTGAAGCGGACCGCCAGGCAATTTCTTC 602  
Qy 481 CGACAGATCGTGTCTGCCCTGCATCTGCGCACACAGAACGGGATCGTTTACCGAGATCTC 540  
Db 603 CGACAGATCGTGTCTGCCCTGCATCTGCGCACACAGAACGGGATCGTTTACCGAGATCTC 662  
Qy 541 AAGCTGGAACACATCTTCTAGATGCCAATGGAAACATCAAGATTGCTGACTTTGGGCTC 600  
Db 663 AAGCTGGAACACATCTTCTAGATGCCAATGGAAACATCAAGATTGCTGACTTTGGGCTC 722  
Qy 601 TCCAACTGTATCCAAAGGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTCTCTACGCC 660  
Db 723 TCCAACTGTATCCAAAGGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTCTCTACGCC 782  
Qy 661 TCGCTCAGATAGTCAAGCGGAGCCCTATGTGGGCCAGAGTGGACGTGTCTCTG 720  
Db 783 TCGCTCAGATAGTCAAGCGGAGCCCTATGTGGGCCAGAGTGGACGTGTCTCTG 842  
Qy 721 GCGCTTCTCTGTATCATCTTGTGATGGACACCATGCCCTTTGACGGGAGGATCATATA 780  
Db 843 GCGCTTCTCTGTATCATCTTGTGATGGACACCATGCCCTTTGACGGGAGGATCATATA 902  
Qy 781 ACATGTGTGAAGCAAAATCAGTAAACGGGCTTACCGTGAAGCGCCCAAGCCGTCCGATGCC 840  
Db 903 ACATGTGTGAAGCAAAATCAGTAAACGGGCTTACCGTGAAGCGCCCAAGCCGTCCGATGCC 962  
Qy 841 TGTGGCTGATCCGGTGGCTGTTAATGTGTAACCCACCCACCGTCCGGCCACACTTGGAGAT 900  
Db 963 TGTGGCTGATCCGGTGGCTGTTAATGTGTAACCCACCCACCGTCCGGCCACACTTGGAGAT 1022  
Qy 901 GTAGCCAGTCAATTTGGTGGGTCAACTGGGGTTTACACCCAGGAGTCGGGGACAGGAAGCC 960  
Db 1023 GTAGCCAGTCAATTTGGTGGGTCAACTGGGGTTTACACCCAGGAGTCGGGGACAGGAAGCC 1082  
Qy 961 CTGGCTGAGGGTGGGACCCCTAGTGTGATCTTTGGCCGGGCTCCATGGCGGACTGTTA 1020  
Db 1083 CTGGCTGAGGGTGGGACCCCTAGTGTGATCTTTGGCCGGGCTCCATGGCGGACTGTTA 1142  
Qy 1021 CGTGGCTCTCTCGGCCCCCTCTGGAGATGGAGCCAAAGTGTGACGCTTCTTCAAGCAG 1080  
Db 1143 CGTGGCTCTCTCGGCCCCCTCTGGAGATGGAGCCAAAGTGTGACGCTTCTTCAAGCAG 1202  
Qy 1081 CAGTGGCGGAGGTGAAGCACTGTACTTGGCTGGAGCGGCAACATTTCTTTAAGAAG 1140  
Db 1203 CAGTGGCGGAGGTGAAGCACTGTACTTGGCTGGAGCGGCAACATTTCTTTAAGAAG 1262  
Qy 1141 TCCGAAAGAGATGATAGTGGTCAAAATCTGCAAGTGAACCCGGCTGAGGATACCTCT 1200  
Db 1263 TCCGAAAGAGATGATAGTGGTCAAAATCTGCAAGTGAACCCGGCTGAGGATACCTCT 1322  
Qy 1201 TCTCGCCCTGGCAAGACAGCCTTAAAGTTCGAAAGGCAATTTCTCAAGAAAGAGTCTCT 1260  
Db 1323 TCTCGCCCTGGCAAGACAGCCTTAAAGTTCGAAAGGCAATTTCTCAAGAAAGAGTCTCT 1382  
Qy 1261 ACTCTGTACGGAGGTACAGGAGGACCTTCAAGAACTCAGAACGGTGTCTGTATCTCA 1320  
Db 1383 ACTCTGTACGGAGGTACAGGAGGACCTTCAAGAACTCAGAACGGTGTCTGTATCTCA 1442  
Qy 1321 GGGCAGCTCTCTCTGTATCTCTCTCCAGGAGGAGGATCTTAAAGATCTCGA 1380  
Db 1443 GGGCAGCTCTCTCTGTATCTCTCTCCAGGAGGAGGATCTTAAAGATCTCGA 1502

Qy 1381 CAGCGTGAATCTGTTACTACTCTCTCCAGAGCCAGCGAGTCTGGGAACTCTTAGAC 1440  
Db 1503 CAGCGTGAATCTGTTACTACTCTCTCCAGAGCCAGCGAGTCTGGGAACTCTTAGAC 1562  
Qy 1441 GCCAGTGAATCTGTTAGTGGGACCCCGTGGAGCAGAGTCTCCACAGGCTTCAGGG 1500  
Db 1563 GCCAGTGAATCTGTTAGTGGGACCCCGTGGAGCAGAGTCTCCACAGGCTTCAGGG 1622  
Qy 1501 CTCTCTCCACCGCAAGGCAATTTCAAACCTCAATGGCAAGTCTCCCGCACAAGCCTTA 1560  
Db 1623 CTCTCTCCACCGCAAGGCAATTTCAAACCTCAATGGCAAGTCTCCCGCACAAGCCTTA 1682  
Qy 1561 GAAGGCACTACCCCTAGACACTTTGGCTCCCTGGAACCACTGGCTCTCCCATCTGCA 1620  
Db 1683 GAAGGCACTACCCCTAGACACTTTGGCTCCCTGGAACCACTGGCTCTCCCATCTGCA 1742  
Qy 1621 GCCCGCCAGCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTGTCTCCGAGTCC 1680  
Db 1743 GCCCGCCAGCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTGTCTCCGAGTCC 1802  
Qy 1681 TTTGACCAATTTGACTTGGCTGAACGCTTTCCCGAAACCCCACTGAGGGGCTGTGTCT 1740  
Db 1803 TTTGACCAATTTGACTTGGCTGAACGCTTTCCCGAAACCCCACTGAGGGGCTGTGTCT 1862  
Qy 1741 GTGGACAACTGAGGGGCTTTGAGGAGCCCTCCCTCAGAGGCTCTGAAGCGATGGTGGCAG 1800  
Db 1863 GTGGACAACTGAGGGGCTTTGAGGAGCCCTCCCTCAGAGGCTCTGAAGCGATGGTGGCAG 1922  
Qy 1801 GAATCTTGGGGATAGCTGCTTTCTTGACAGACTGCCAAGAGTGTACTGCAGCTAC 1860  
Db 1923 GAATCTTGGGGATAGCTGCTTTCTTGACAGACTGCCAAGAGTGTACTGCAGCTAC 1982  
Qy 1861 AGACAAGCCCTAGGAATCTGCTCAAAGCTCAGC 1893  
Db 1983 AGACAAGCCCTAGGAATCTGCTCAAAGCTCAGC 2015

## RESULT 2

AX381046  
LOCUS AX381046 2027 bp DNA linear PAT 18-MAR-2002  
DEFINITION Sequence 87 from Patent WO212456.  
ACCESSION AX381046  
VERSION AX381046.1 GI:19575868  
KEYWORDS  
SOURCE Mus sp.  
ORGANISM Mus sp.  
REFERENCE 1  
AUTHORS Drucker, D.J., Rosen, C.F. and Lefebvre, D.L.  
TITLE Ampk-related serine/threonine kinase, designated snark  
JOURNAL Patent: WO 0212456-A 87 14-FEB-2002;  
1149336 ONTARIO INC. (CA)  
FEATURES  
source  
1. .2027  
/organism="Mus sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10095"

## ORIGIN

Query Match 99.7%; Score 1888.2; DB 6; Length 2027;  
Best Local Similarity 99.8%; Pred. No 0;  
Matches 1890; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 ATGGAGTCCGGTGGCTTACTCCAGGCCCGAGGCTCCCTCGGCTCCGCTGGCC 60  
Db 75 ATGGAGTCCGGTGGCTTACTCCAGGCCCGAGGCTCCCTCGGCTCCGCTGGCC 134  
Qy 61 TCGGAGAGCGCCGCGCTGGGAGCGGCTCATCAAGTCCGCTTAAACCTCTGATGAG 120  
Db 135 TCGGAGAGCGCCGCGCTGGGAGCGGCTCATCAAGTCCGCTTAAACCTCTGATGAG 194

121 AAGCAGCGGTGAAGCGGACCCATCAAAACAACTGGGCGCCCGCTACGAGTTCTCTG 180  
195 AAGCAGCGGTGAAGCGGACCCATCAAAACAACTGGGCGCCCGCTACGAGTTCTCTG 254  
181 GAGACGCTGGGCAAGGCGACCTACGGGAAGGTGAAGAACGACGAGAGCTCGGGCGGT 240  
255 GAGACGCTGGGCAAGGCGACCTACGGGAAGGTGAAGAACGACGAGAGCTCGGGCGGT 314  
241 CTGGTGGCCATCAAGTCCATCAGGAAAGACAAATCAAGATGAGCAGGATCTCTGCAC 300  
315 CTGGTGGCCATCAAGTCCATCAGGAAAGACAAATCAAGATGAGCAGGATCTCTGCAC 374  
301 ATACGAGGAGATGAGATCATGCTTCACTCAACCAACCCCAATCAATGCGCATCCAT 360  
375 ATAAGGAGGAGATCGAGATCATGCTTCACTCAACCAACCCCAATCAATGCGCATCCAT 434  
361 GAGTGTTCAGATAGCAGCAGATTTGATTTGATGATGATGATGATGATGATGATGAT 420  
435 GAGTGTTCAGATAGCAGCAGATTTGATTTGATGATGATGATGATGATGATGATGAT 494  
421 CTGTATGATTACATCAGTGAAGCGGACCGGCTGAGTGAGCGGAGCGCAGGCAATTTCTTC 480  
495 CTGTATGATTACATCAGTGAAGCGGACCGGCTGAGTGAGCGGAGCGCAGGCAATTTCTTC 554  
481 CGACAGATCGTGTCTGCCCTGCACTACTGCGCACAGAACCGGATCGTTTCAACCGAGATCTC 540  
555 CGACAGATCGTGTCTGCCCTGCACTACTGCGCACAGAACCGGATCGTTTCAACCGAGATCTC 614  
541 AAGCTGAAAAACATCCTTTAGATGCCAATGGAAACATCAAGATTTGCTGATTTGGGCTC 600  
615 AAGCTGAAAAACATCCTTTAGATGCCAATGGAAACATCAAGATTTGCTGATTTGGGCTC 674  
601 TCCAACTGTACCAACAAAGCAAGTCTCTCAGACGTTCTGTGGAGCCCTCTCTAGCGCC 660  
675 TCCAACTGTACCAACAAAGCAAGTCTCTCAGACGTTCTGTGGAGCCCTCTCTAGCGCC 734  
661 TCGCCTGAGATAGTCAACGGGAGCCCTATGTGGGCCCGAGAGTGGAAGCTGTCTCTG 720  
735 TCGCCTGAGATAGTCAACGGGAGCCCTATGTGGGCCCGAGAGTGGAAGCTGTCTCTG 794  
721 GCGGTTCTCTGTACATCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 780  
795 GCGGTTCTCTGTACATCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 854  
781 ACATGCTGAAGCAAAATCAGTAAACGGGCTTACCGTGAGCGGCCCAAGCGCTCGATGCC 840  
855 ACATGCTGAAGCAAAATCAGTAAACGGGCTTACCGTGAGCGGCCCAAGCGCTCGATGCC 914  
841 TGTGGCCTGATCCGGTGGCTGTTAATGTTGAACCCCAACCGCTCGGGCCACACTGGAGGAT 900  
915 TGTGGCCTGATCCGGTGGCTGTTAATGTTGAACCCCAACCGCTCGGGCCACACTGGAGGAT 974  
901 GTAGCCAGTCAATTGGTGGTCACTGGGGTTACACCAACCGAGTGGGGACAGGAAGCC 960  
975 GTAGCCAGTCAATTGGTGGTCACTGGGGTTACACCAACCGAGTGGGGACAGGAAGCC 1034  
961 CTGGCTGAGGCTGGGCAACCTAGTGGGCTTTGGCGGCGCTTCAATGGCGGACTGGTTA 1020  
1035 CTGGCTGAGGCTGGGCAACCTAGTGGGCTTTGGCGGCGCTTCAATGGCGGACTGGTTA 1094  
1021 CGTGGCTCTCGCGCCCTCTCGGAGAAATGGAGCCCAAGGTGTGAGCTTCTTCAAGCAG 1080  
1095 CGTGGCTCTCGCGCCCTCTCGGAGAAATGGAGCCCAAGGTGTGAGCTTCTTCAAGCAG 1154  
1081 CAGTGGCGGAGGTGAAGCACTGTACCTGGGCTGGAGCGGCAATTTCTTTAAGAAG 1140  
1155 CAGTGGCGGAGGTGAAGCACTGTACCTGGGCTGGAGCGGCAATTTCTTTAAGAAG 1214  
1141 TCCGGAAGGAGATGATGAGTCAATGCTCAAGGTGACCCGGCTGAGGATACCTCT 1200  
1215 TCCGGAAGGAGATGATGAGTCAATGCTCAAGGTGACCCGGCTGAGGATACCTCT 1274  
1201 TCTCGCCCTGGCAAGAGCAGCTTTAAGCTTTCCGAAAGGCATTTCTCAAGAAAAAGTCTCT 1260

1275 TCTCGCCCTGGCAAGAGCAGCTTAAAGTTTCGAAAGGCATTTCTCAAGAAAAAGTCTCT 1334  
1261 ACCTCGTCAAGGAGGTACAGGAGGACCTCAGNACTCAGACCGGTCCCTGATCTCCA 1320  
1335 ACCTCGTCAAGGAGGTACAGGAGGACCTCAGNACTCAGACCGGTCCCTGATCTCCA 1394  
1321 GGGCAGCCTCTCCCTGCTGTATCCCTGCTCCCAAGGAAAGGCATCTTAAAGAGTCTCGA 1380  
1395 GGGCAGCCTCTCCCTGCTGTATCCCTGCTCCCAAGGAAAGGCATCTTAAAGAGTCTCGA 1454  
1381 CAGGTGAATCTGGTTACTACTCTCTCCAGAGCCAGAGTCTGGGAACTCTTAGAC 1440  
1455 CAGGTGAATCTGGTTACTACTCTCTCCAGAGCCAGAGTCTGGGAACTCTTAGAC 1514  
1441 GCCAGTGAATCTGGTTAGTGGGACCCGCTGGAGCAGAGTCTCCACAGGCTTCAGGG 1500  
1515 GCCAGTGAATCTGGTTAGTGGGACCCGCTGGAGCAGAGTCTCCACAGGCTTCAGGG 1574  
1501 CTCTCTCTCCAGCGCAAGGCAATTTCAAACTCAATGGCAAGTTCTCCCGCACAGCCTTA 1560  
1575 CTCTCTCTCCAGCGCAAGGCAATTTCAAACTCAATGGCAAGTTCTCCCGCACAGCCTTA 1634  
1561 GAAGGCACTACCTCTAGCACCCTTTGGCTCCCTGGACCAACTGGCCTCTCCATCTGCA 1620  
1635 GAAGGCACTACCTCTAGCACCCTTTGGCTCCCTGGACCAACTGGCCTCTCCATCTGCA 1694  
1621 GCCGGGCCAGCGCCCTCAGGGGCTGTAGTGAGGACAGCATCTCTCTCCGAGTCC 1680  
1695 GCCGGGCCAGCGCCCTCAGGGGCTGTAGTGAGGACAGCATCTCTCTCCGAGTCC 1754  
1681 TTTGACCAATTTGGACTTTGGCTGAAAGCTTTCCGAAACCCCACTGAGGGGCTGTGTCT 1740  
1755 TTTGACCAATTTGGACTTTGGCTGAAAGCTTTCCGAAACCCCACTGAGGGGCTGTGTCT 1814  
1741 GTGGAACAACCTGAGGGGCTTTGAGCAGCCTCTCCCTCAGAAAGGTCTGAAGCAGATGGTGCAG 1800  
1815 GTGGAACAACCTGAGGGGCTTTGAGCAGCCTCTCCCTCAGAAAGGTCTGAAGCAGATGGTGCAG 1874  
1801 GAATCTCTGGGGATAGTCTCTCTCAGACAGCTGCGCAAGGAGTCTGAGCAGCTAC 1860  
1875 GAATCTCTGGGGATAGTCTCTCTCAGACAGCTGCGCAAGGAGTCTGAGCAGCTAC 1934  
1861 AGACAAGCCCTAGGAATCTCTCAAAAGCTCAGC 1893  
1935 AGACAAGCCCTAGGAATCTCTCAAAAGCTCAGC 1967

## RESULT 3

BC033302

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BC033302 2917 bp mRNA linear ROD 04-NOV-2003  
Mus musculus RIKEN cDNA 1200013B22 gene, mRNA (cDNA clone MGC:30598  
IMAGE:3663595), complete cds.

BC033302

BC033302.1 GI:23271085

MGC.

Mus musculus (house mouse)

Mus musculus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 2917)

Klausner, R.D., Collins, F.S., Wagner, L.H., Derge, J.G.,

Staub, R.D., Collins, F.S., Wagner, L.H., Derge, J.G.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F.,

Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshyuk, S.,

Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bufford, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dicksen, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalhus, D.E., Schnerich, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2388257

PUBMED

REFERENCE

2 (bases 1 to 2917)

Strausberg, R.

Direct Submission

Submitted (26-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)

Gunaratne, P. H., Garcia, A. M., Lu, X., Hulyk, S. W., Loulseghe, H., Kowis, C. R., Sneed, A. J., Martin, R. G., Muzny, D. M., Nanavati, A. N., Gibbs, R. A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 40 Row: d Column: 21

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27229181.

## FEATURES

Location/Qualifiers

1..2917

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="Czech II"

/db\_xref="taxon:10090"

/clone="MGC:30598 IMAGE:3663595"

/tissue\_type="Mammary tumor metastasized to lung."

MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR enhancer.

/clone\_lib="NCI CGAP Lu30"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

1..2917

/gene="1200013B22Rik"

/note="synonym: Shark"

/db\_xref="LocusID:74137"

/db\_xref="MGI:1921387"

81..1976

/gene="1200013B22Rik"

/codon\_start=1

/product="1200013B22Rik protein"

/protein\_id="AAH33302.1"

/db\_xref="GI:23271086"

/db\_xref="LocusID:74137"

/db\_xref="MGI:1921387"

/translation="MESVALLQPSQASALASASARPLADGLIKSPKPLMKQAV KRHHKHLNHRFEFLTGKTYGKVKARESSGRILVAIKSIRKDKIKDEODLLHIR REIEMSLNPHI IAHIEVFENSKIIVIMYASRGDLYDIYSERPLSERDARHFF ROIVSALHYCHQNGI VHRDLKENLIDANGNIKIADFGLSNLYHRKFLQTPCGSPL YASPEIVNGPYPVPEVDSVLSGLVLYLVHGTMPDPGDHKLTVQISNGAVREPPK PSDACGLRWLLWNPTRATLEDVASHWVWNGYTTGVGEQALREGGHPSGDSGRA SMADWLRRSRPILENGAKVCSFFKQHPVGGGTVFGLERQELKSKRKNENQALIQ GPADDTSSRPFSGSSLUKPKGLIKKSSSTSSGVEDPQELRPVPTPGQVPVAVSLI PRKGILKRSQRESGYSSPEPSESSELLDASDVFSVSGDPVEQKSPQASGLLLHRKI

## gene

## CDS

misc\_feature

1200013B22Rik"

/note="S.TK; Region: Serine/Threonine protein kinases, catalytic domain. Phosphotransferases of the serine or threonine-specific kinase subfamily. The enzymatic activity of these protein kinases is controlled by phosphorylation of specific residues in the activation segment of the catalytic domain, sometimes combined with reversible conformational changes in the C-terminal autoregulatory tail"

/db\_xref="CDD:cd00180"

ORIGIN

Query Match 98.9%; Score 1872.2; DB 10; Length 2917;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 1880; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 ATGGAGTCGGTGGCCCTTACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCCGCCCTGGCC 60

Db 81 ATGGAGTCGGTGGCCCTTACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCCGCCCTGGCC 140

Qy 61 TCGGAGAGCGCCCGCGGCTGGCGGACGGGCTCATCAAGTCGCCTAAACCTCTGATGAAG 120

Db 141 TCGGAGAGCGCCCGCGGCTGGCGGACGGGCTCATCAAGTCGCCTAAACCTCTGATGAAG 200

Qy 121 AAGCAGCGGTGAAGCGGCACCATCAAAACACAACTCGCGGCACCGCTACGAGTTCCTG 180

Db 201 AAGCAGCGGTGAAGCGGCACCATCAAAACACAACTCGCGGCACCGCTACGAGTTCCTG 260

Qy 181 GAGACCTGGGCAAGGGCACCCTACGGGAAGGTGAAGAAGCAGAGAGCTCGGGCGT 240

Db 261 GAGACCTGGGCAAGGGCACCCTACGGGAAGGTGAAGAAGCAGAGAGCTCGGGCGT 320

Qy 241 CTGGTGGCCATCAAGTCCATCAGAAAGACAAATCAAGATCAGCAGAGATCTGCTGCAC 300

Db 321 CTGGTGGCCATCAAGTCCATCAGAAAGACAAATCAAGATCAGCAGAGATCTGCTGCAC 380

Qy 301 ATACGAGGAGAGATTGAGATCATGTCTTCACTCAACCAACCCCAATATTCGATCCAT 360

Db 381 ATACGAGGAGAGATTGAGATCATGTCTTCACTCAACCAACCCCAATATTCGATCCAT 440

Qy 361 GAAGTCTTTGAGNATAGCAGCAGATGTGTGTTGTCATGGAGTATCCAGCCGAGCGAT 420

Db 441 GAAGTCTTTGAGNATAGCAGCAGATGTGTGTTGTCATGGAGTATCCAGCCGAGCGAT 500

Qy 421 CTGTATGATTACATCAGTCAGCGGCCACCGCTGAGTGAGCGGACCGCAGGCAATTTCTTC 480

Db 501 CTGTATGATTACATCAGTCAGCGGCCACCGCTGAGTGAGCGGACCGCAGGCAATTTCTTC 560

Qy 481 CGACAGATCGTGTCTGCCCTGCACTACTGCCACAGAACGGGATCTTCAACGAGATCTC 540

Db 561 CGACAGATCGTGTCTGCCCTGCACTACTGCCACAGAACGGGATCTTCAACGAGATCTC 620

Qy 541 AAGCTCGAAAACATCTTCTAGATGCCAATGGAAACATCAAGATTCGTGCTTTGGCCTC 600

Db 621 AAGCTCGAAAACATCTTCTAGATGCCAATGGAAACATCAAGATTCGTGCTTTGGCCTC 680

Qy 601 TCCAACTGTACCAAAAGCAAGTTCCTCCAGACGCTTCTGTGGAGCCCTCTCTACGCC 660

Db 681 TCCAACTGTACCAAAAGCAAGTTCCTCCAGACGCTTCTGTGGAGCCCTCTCTACGCC 740

Qy 661 TGCCTTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCAAGAGTGAGAGTGTCTCTG 720

Db 741 TGCCTTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCAAGAGTGAGAGTGTCTCTG 800

Qy 721 GCGTCTTCTCTGATACATCTGTGTGATGCACCATGCTCCCTTTGACGGGAGGATCAATA 780

Db 801 GCGTCTTCTCTGATACATCTGTGTGATGCACCATGCTCCCTTTGACGGGAGGATCAATA 860

Qy 781 AACTGTGAAGCAATCAAGTAAACGGGGCTTACCGTGAGCCGCCCAAGCCGTCCTCGATCC 840

Db 861 AACTGTGTGAAGCAAAATCAGTAACGGGGCTTTACCGTGAGCGGCCCAAGCCGTCGCATGCC 920  
QY 841 TGTGGCCTGATCCGGTGGCTGTTAATGTGAACCCCAACCGTCCGGCCACACTGGAGAT 900  
Db 921 TGTGGCCTGATCCGGTGGCTGTTAATGTGAACCCCAACCGTCCGGCCACACTGGAGAT 980  
QY 901 GTAGCCAGTCAATTGGTGGGTCAACTGGGGTTTACACCAACCGAGTCCGGGAACAGGAAGCC 960  
Db 981 GTAGCCAGTCAATTGGTGGGTCAACTGGGGTTTACACCAACCGAGTCCGGGAACAGGAAGCC 1040  
QY 961 CTGGCTGAGGTGGGACCCCTAGTGTGTGACTTTTGGCCGGGCTTCATGCGCGACTGGTTA 1020  
Db 1041 CTGGCTGAGGTGGGACCCCTAGTGTGTGACTTTTGGCCGGGCTTCATGCGCGACTGGTTA 1100  
QY 1021 CGTCGCTCTCGCGCCCTCTCGAGAAATGGAGCCAGGTGTGAGCTTCTTCAAGCAG 1080  
Db 1101 CGTCGCTCTCGCGCCCTCTCGAGAAATGGAGCCAGGTGTGAGCTTCTTCAAGCAG 1160  
QY 1081 CACGTGCGGGAGGTGGAAGCACTGTACTCGGGCTGGAGCGGCAACATTTCTTTAAGAAG 1140  
Db 1161 CACGTGCGGGAGGTGGAAGCACTGTACTCGGGCTGGAGCGGCAACATTTCTTTAAGAAG 1220  
QY 1141 TCCGGAAGGAGATGACATCGCTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCTCT 1200  
Db 1221 TCCGGAAGGAGATGACATCGCTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCTCT 1280  
QY 1201 TCTCGCTCTGCAAGAGCGCTTAAGTCTCGAAGGCAATCTCAAGAAAAGTCTCT 1260  
Db 1281 TCTCGCTCTGCAAGAGCGCTTAAGTCTCGAAGGCAATCTCAAGAAAAGTCTCT 1340  
QY 1261 ACCTCGTCAGGGAGGTACAGAGGACCCCTCAGGAATCTCAGACCGGTGCTGTACTCCA 1320  
Db 1341 ACCTCTTCAGGGAGGTACAGAGGACCCCTCAGGAATCTCAGACCGGTGCTGTACTCCA 1400  
QY 1321 GGGCAGCGCTGCTCTGTGTATCTCTCTCCCAAGGAAAGGCATCTTTAAGAACTCTGA 1380  
Db 1401 GGGCAGCGCTGCTCTGTGTATCTCTCTCCCAAGGAAAGGCATCTTTAAGAACTCTGA 1460  
QY 1381 CAGCGTGAATCTGTACTACTCTCTCCAGAGCCAGCGAGTCTGGGAACTCTTAGAC 1440  
Db 1461 CAGCGTGAATCTGTACTACTCTCTCCAGAGCCAGCGAGTCTGGGAACTCTTAGAC 1520  
QY 1441 GCCAGTGTATGTTGTGAGTGGGACCCCTGAGGACAGAGTCTCAGACGGCTTCAGGG 1500  
Db 1521 GCCAGTGTATGTTGTGAGTGGGACCCCTGAGGACAGAGTCTCAGACGGCTTCAGGG 1580  
QY 1501 CTCCTCTCTCCACGCAAGGGCAATCTCAAACTCAATGCAAGTCTCTCCGCAAGCCCTTA 1560  
Db 1581 CTCCTCTCTCCACGCAAGGGCAATCTCAAACTCAATGCAAGTCTCTCCGCAAGCCCTTA 1640  
QY 1561 GAAGGCACACTCCCTAGGACCTTTGGCTCTCTGACCAACTGGGCTCTCCCATCTCTGCA 1620  
Db 1641 GAAGGCACACTCCCTAGGACCTTTGGCTCTCTGACCAACTGGGCTCTCCCATCTCTGCA 1700  
QY 1621 GCCGGGCCACCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTGTCTCTCCGAGTCC 1680  
Db 1701 GCCGGGCCACCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTGTCTCTCCGAGTCC 1760  
QY 1681 TTTGACCAATTTGAGTCTGCTGAAAGTCTTCCCGAAACCCCACTGAGGGGCTGTGTGTCT 1740  
Db 1761 TTTGACCAATTTGAGTCTGCTGAAAGTCTTCCCGAAACCCCACTGAGGGGCTGTGTGTCT 1820  
QY 1741 GTGACAACTGAGGGGCTTGAAGAGCTCTCTCAGAGTCTGAACCGATGTGGGAG 1800  
Db 1821 GTGACAACTGAGGGGCTTGAAGAGCTCTCTCAGAGTCTGAACCGATGTGGGAG 1880  
QY 1801 GAATCTTTGGGGATAGCTGCTTTTCTCTGACAGACTGCCAAGAGGTGACTGCAGCCTAC 1860  
Db 1881 GAATCTTTGGGGATAGCTGCTTTTCTCTGACAGACTGCCAAGAGGTGACTGCAGCCTAC 1940  
QY 1861 AGACAAGCCCTAGGAATCTGCTCAAAAGCTCAGC 1893

1941 AGACAAGCCCTAGGAATCTGCTCGAAGCTCAGC 1973  
Db 1941 AGACAAGCCCTAGGAATCTGCTCGAAGCTCAGC 1973  
RESULT 4  
BC046833  
LOCUS  
DEFINITION  
Mus musculus RIKEN cDNA 120013B22 gene, mRNA (cDNA clone MGC:61244 IMAGE:6816569), complete cds.  
ACCESSION  
BC046833  
VERSION  
BC046833.1  
KEYWORDS  
GI:29145027  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
REFERENCE  
1 (bases 1 to 2938)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myer, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smallos, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
TITLE  
human and mouse cDNA sequences  
JOURNAL  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
MEDLINE  
22388257  
PUBMED  
12477932  
REFERENCE  
2 (bases 1 to 2938)  
AUTHORS  
Strausberg, R.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (13-FEB-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-x@mail.nih.gov](mailto:cgapbs-x@mail.nih.gov)  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library Preparation: M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.  
Web site: <http://genome.uiowa.edu>  
Contact: [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fisher, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: Place: Row: Column: 0  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27229181.  
FEATURES  
Location/Qualifiers  
1..2938  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="MGC:61244 IMAGE:6816569"

```
/tissue type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"
/clone_lib="NIH_BMAP_FW0"
/lab_host="DH10B"
/notes="vector: pYX-ASC"
1. .2938
/gene="1200013B22Rik"
/notes="synonym: Shark"
/db_xref="LocustID:74137"
/db_xref="MGI:1921387"
81. .2000
/gene="1200013B22Rik"
/codon_start=1
/product="1200013B22Rik protein"
/protein_id="AAH46833.1"
/db_xref="GI:29145028"
/db_xref="LocustID:74137"
/db_xref="MGI:1921387"
/translacion="MESVALIQRPSALASASARPLADGLIKSPKPLMKQAV
KRHHKHLNRHYEFLTLGKTYGVKKARESSGRVIAIKSIRKDKI KBOODLHLR
REIIMSLNPHPIIAIHEVGRSLVTVFENSSKIIVMEVASRGLDYVISEPRLS
ERDARHFRQIVSALHYCHQNGIVHRDLKLENILLDANGNIKIADFGLSNLYHKXEL
QTCFGLSPASPELVNKPYPVGPEDVDSWSLGLVLIIVHGTMPDGDHKLTVKQISN
GAYREPKPDADCLIRLWLVNPTRRATLEDVASHWVNWGYTTGVGEQALREGGH
PSGDFGRADWLRSSRLLENKAVKVCFFKQHPVGGSTVPLGRQRLSLKSRKE
NDMAQNLQGGPAEDTSRPGKSLKPKGILKKKSSTSSGEVQBDPDLRPVDPDTCQ
PVPAVSLPRKILKRSORESGYSSPEPSESGELLDADVFVSGDPVQKSPQASG
LLHRKGLILKNGKESFALGTPSTFGLDQLASSHPAARPERPRGAVSESLISG
ESFDQLDLPERLPPTARGVSVSNLRGLGEPSEGLKRWQBSLGSDFSLTDCQEV
TAAYRQALGIGCSKLS"
249. .1025
/misc_feature
/gene="1200013B22Rik"
/notes="S_TKc; Region: Serine/Threonine protein kinases,
catalytic domain. Phosphotransferases of the serine or
threonine-specific kinase subfamily. The enzymatic
activity of these protein kinases is controlled by
phosphorylation of specific residues in the activation
segment of the catalytic domain, sometimes combined with
reversible conformational changes in the C-terminal
autoregulatory tail"
/db_xref="CDD:cd00180"
ORIGIN
Query Match 98.2%; Score 1859; DB 10; Length 2938;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 0; Indels 24; Gaps 1;
Qy 1 ATGGAGTCGGTGGCTTACTACAGCGCCGAGCAGGCTCCCTCGGCTCGCCCTCGCCCTGGCC 60
Db 81 ATGGAGTCGGTGGCTTACTACAGCGCCGAGCAGGCTCCCTCGGCTCGCCCTCGCCCTGGCC 140
Qy 61 TCGGAGAGCGCCGCGCGCTGGCGGACGGGCTCATCAAGTCGCCTAAACCTCTCATGAAG 120
Db 141 TCGGAGAGCGCCGCGCGCTGGCGGACGGGCTCATCAAGTCGCCTAAACCTCTCATGAAG 200
Qy 121 AAGCAGCGGTGAAGCGGCACCATCAAAAACAACCTCGGCGCACCGCTACGAGTTCCTG 180
Db 201 AAGCAGCGGTGAAGCGGCACCATCAAAAACAACCTCGGCGCACCGCTACGAGTTCCTG 260
Qy 181 GAGACGCTGGGCAAGGCACTACGGGAAGGTGAAGAGGACAGAGAGCTCGGGCGGT 240
Db 261 GAGACGCTGGGCAAGGCACTACGGGAAGGTGAAGAGGACAGAGAGCTCGGGCGGT 320
Qy 241 CTGTGGCCATCAAGTCCATCAGAAAGACAAAATCAAGATGACGAGGATCTCTGCAC 300
Db 321 CTGTGGCCATCAAGTCCATCAGAAAGACAAAATCAAGATGACGAGGATCTCTGCAC 380
Qy 301 ATACGGAGGAGATTGAGATCATGTCTTCACTCAACCCACCCCAACCATTCGCCATCAT 360
Db 381 ATACGGAGGAGATTGAGATCATGTCTTCACTCAACCCACCCCAACCATTCGCCATCAT 440
Qy 361 GAA-----GTGTTTGAATAGCAGCAGATGTGATGTC 396
Db 441 GAAATGGGCGAGATCACGCTTAGTGACTGTGTGTTTGAATAGCAGCAGATGTGATGTC 500
```



Qy 1477 CAGAAGTCTCCACAGGCTTCAGGGCTCTCTCTCCACCGCAAGGCAATCTCAAACTCAAT 1536  
Db CAGAAGTCTCCACAGGCTTCAGGGCTCTCTCTCCACCGCAAGGCAATCTCAAACTCAAT 1640  
Qy 1537 GGCAAGTCTCCCGCAGAGCCTTAGAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGAC 1596  
Db GGCAAGTCTCCCGCAGAGCCTTAGAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGAC 1700  
Qy 1597 CAACTGGCTCTCTCCCATCTCTGACGCCGGGCCAGCGCCCTCAGGGGCTGTGAGTGAG 1656  
Db CAACTGGCTCTCTCCCATCTCTGACGCCGGGCCAGCGCCCTCAGGGGCTGTGAGTGAG 1760  
Qy 1657 GACAGCATCTCTCGAGTCTCTTGAACCAATTGAGCTTTCGTCGAACTGTTCCCGAA 1716  
Db GACAGCATCTCTCGAGTCTCTTGAACCAATTGAGCTTTCGTCGAACTGTTCCCGAA 1820  
Qy 1717 ACCCCACTGAGGGGCTGTGTCTGTGAGCAACCTGAGGGGCTTGAAGAGCTCTCCCTCA 1776  
Db ACCCCACTGAGGGGCTGTGTCTGTGAGCAACCTGAGGGGCTTGAAGAGCTCTCCCTCA 1880  
Qy 1777 GAAGGTCTGAAGCGATGGTGGCAGGAATCTTGGGGGATAGTCTCTTCTCTGACAGAC 1836  
Db GAAGGTCTGAAGCGATGGTGGCAGGAATCTTGGGGGATAGTCTCTTCTCTGACAGAC 1940  
Qy 1837 TGCCAAAGAGTGACTGAGGCTTACAGACAGCCCTAGGAATCTGCTCAAAAGCTCAGC 1893  
Db TGCCAAAGAGTGACTGAGGCTTACAGACAGCCCTAGGAATCTGCTCAAAAGCTCAGC 1997

RESULT 5  
AX380986  
LOCUS AX380986 2026 bp DNA linear PAT 18-MAR-2002  
DEFINITION Sequence 27 from Patent WO0212456.  
ACCESSION AX380986  
VERSION AX380986.1 GI:19575826  
KEYWORDS  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1  
AUTHORS Drucker, D.J., Rosen, C.F. and Lefebvre, D.L.  
TITLE Ampk-related serine/threonine kinase, designated snark  
JOURNAL Patent: WO 0212456-A 27 14-FEB-2002;  
1149336 ONTARIO INC. (CA)  
FEATURES  
source location/Qualifiers  
1..2026  
/organism="Rattus sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10118"

ORIGIN

Query Match 88.9%; Score 1683.4; DB 6; Length 2026;  
Best Local Similarity 93.4%; Pred. No. 0;  
Matches 1769; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

Qy 1 ATGAGTTCGGTGGCTTACTCCAGCGCCGAGCAGGCTCCCTCGGCTCGGCTCGGCC 60  
Db ATGAGTTCGGTGGCTTACTACCGCGCGGGAACCTGGCTCCCTCGGCTCGGCTCGGCC 134  
Qy 61 TCGGAGAGCGCCCGCGCTGGCGGACGGGCTCATCAAGTCGCCCTAAACCTCTGATGAAG 120  
Db ACGGAGAGCGCCCGCGCTGGCGGACGGGCTCATCAAGTCGCCCTAAACCTCTGATGAAG 194  
Qy 121 AAGCAGCGGTGAAGCGGCAACCATCAAAACAACTCTGGGCAACCGCTACGAGTTCCTG 180  
Db AAGCAGCGGTGAAGCGGCAACCATCAAAACAACTCTGGGCAACCGCTACGAGTTCCTG 254  
Qy 181 GAGACGCTGGCAGGGCACTTACGGGAGGTGAAGAGGACACGAGAGCTCGGGCGT 240  
Db GAGACGCTGGCAGGGCACTTACGGGAGGTGAAGAGGACACGAGAGCTCGGGAGC 314

Qy 241 CTGTGGCCATCAAGTCCATCAGGAAAGACAAAATCAAAGATGACGAGGATCTGCTGCAC 300  
Db CTGTGGCCATCAAGTCTATCAGGAAGGACAAAATCAAAGATGACGAGGATCTGTTGCAC 374  
Qy 301 ATACGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAACATCATTTGCCATCAT 360  
Db ATAAGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAACATCATTTGCCATCAT 434  
Qy 361 GAAGTGTTCGAGAAATAGCAGCAAGATTGTGATGTGTATGAGTATGCGCCGAGGGGAT 420  
Db GAAGTGTTCGAGAAATAGCAGCAAGATTGTGATGTGTATGAGTATGCGCCGAGGGGAT 494  
Qy 421 CTGTATGATTAATCACTAGTGAAGCGCCACGGCTGAGTGAAGCGGAGCGCCAGGCAATTTCTTC 480  
Db CTGTATGATTAATCACTAGTGAAGCGCCACGGCTGAGTGAAGCGGAGCGCCAGGCAATTTCTTC 554  
Qy 481 CGACAGATCTGTCTGCTCCCTGCACTACTGCGCAACAGAACGGGATCGTTTACCGAGATCTC 540  
Db CGACAGATCTGTCTGCTCCCTGCACTACTGCGCAACAGAACGGGATCGTTTACCGGAGCCTC 614  
Qy 541 AAGCTGGAAGAAACATCTCTTAGATGCCAATGGAACATCAAGATTGCTGATCTTTGGGCTC 600  
Db AAGCTGGAAGAAACATCTCTTAGATGCCAATGGAACATCAAGATTGCTGATCTTTGGGCTC 674  
Qy 601 TCCAACCTGTATCCAAAGGCAAGTTCTCTCAGACGTTTCTGTGGGAGCCCTCTCTACGCC 660  
Db TCCAACCTGTATCCAAAGGCAAGTTCTCTCAGACGTTTCTGTGGGAGCCCTCTCTATGCC 734  
Qy 661 TCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCACAGAGTGGACAGTGTCTCTGT 720  
Db TCACCTGAGATCGTCAACGGGAAGCCCTATGTGGGCCACAGAGTGGACAGTGTCTCTGT 794  
Qy 721 GGGCTTCTCTGTATACATCTCTGTGCAATGCGCACCATGCTTTGACGGGAGGATCATATA 780  
Db GGGCTTCTCTGTATACATCTCTGTGCAATGCGCACCATGCTTTGACGGGAGGATCATATA 854  
Qy 781 ACATGCTGAAGCAAAATCAGTAACGGGCTTACCGTGAGCGCCCAAGCGCTCCGATGCC 840  
Db ACCTGCTGAAGCAAAATCAGTAACGGGCTTACCGTGAGCGCTTACCGAGAGCGCTCGTATGCC 914  
Qy 841 TGTGGCTGATTCGGTGGCTGTTAATGGTGAAACCCACCGCTGGGGCACACTGGAGGAT 900  
Db TGTGGCTGATTCGGTGGCTGTTAATGGTGAAACCCACCGCTGGGGCACACTCTGGAGGAT 974  
Qy 901 GTAGCCAGTCAATTTGGTGGTCAACTGGGTTTACACCAACCGAGTTCGGGGAGGAGAGCC 960  
Db GTAGCCAGTCAATTTGGTGGTCAACTGGGTTTACACCAACCGAGTTCGGGGAGGAGAGCT 1034  
Qy 961 CTGCGTGAGGCTGGGCAACCTAGTGTGTGACTTTTGGCGGGCTCCATGCGGGAGTGGTTA 1020  
Db CTGGGAGAGGCTGGGCAACCTAGCGGTGACTCTGGCGGGCTCTATGGCGGACTGGTTA 1094  
Qy 1021 GTGCGTCTCTGGCGCCCTCTTGGAGAAATGAGAGCCAAAGTGTGTGACGCTTCTTCAAGCAG 1080  
Db CGTGTCTCTCTGGCGCCCTCTTGGAGAAATGAGAGCCAAAGTGTGTGACGCTTCTTCAAGCAG 1154  
Qy 1081 CAGTGTCCGGAGGTGGAAGCACTGTACCTGGGCTGGAGCGGCAACATCTCTTAAAGAG 1140  
Db CATGTGTCCGGAGGTGGAAGCACTGTACCTGGGCTGGAGCGGCAACATCTCTTAAAGAG 1214  
Qy 1141 TCCGGAAGGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCTCT 1200  
Db TCCGGAAGGAGAAATGACATGGCTCAGACTCTGCAAGTATGACCAAGTGAAGTACTTCC 1274  
Qy 1201 TCTGCGCTCTGGCAAGAGCAGCTTAAAGCTTCGAAAGGCAATTTCTCAAGAAAGAGTCTCT 1260  
Db TCTGCGCTCTGGCAAGAGCAGCTTAAAGCTTCGAAAGGCAATTTCTCTCAAGAAAGAGTCTCT 1334  
Qy 1261 ACCTGTCTGAGGGAGGTACAGGAGGAGCCCTCAGGAACTCAGACGGTGCCTGATCTCCA 1320  
Db CCCTCATCTGGGGAGGTACAGGAGGAGCCCTCAGGAACTCAGACCGAGTGTCAATACCCCA 1394  
Qy 1321 GGGCAGCCTGTCTGCTGTATCCCTGCTCCCAAGGAAAGGCAATCTTAAAGAGTCTCTCGA 1380



```

Db      1395  GGGCAGCTGTCTCTATACCTCTCTCCAGGAGGCGATCTTAAAGAGTCTCGG 1454
Qy      1381  CAGCGTGAATCTGGTTACTCTCTCCAGAGCCAGCGAGTCTGGGAACTCTTAGAC 1440
Db      1455  CAGCGTGAATCTGGTTACTCTCTCCAGAGCCAGTGAAGTCTGGGAACTCTTAGAC 1514
Qy      1441  GCAGTGAATCTGGTTAGTGGGAGCCCGGTGAGCAGAGTCTCCACAGGCTTCAGGG 1500
Db      1515  GCAGTGAATCTGGTTAGTGGGAGCCCGGTGAGCAGAGTCTCCACAGGCTTCAGGG 1574
Qy      1501  CTCCTCTCCACCAAGGCAATCTCAAACTCAATGGCAAGTCTCTCCGCAACAGCTTA 1560
Db      1575  ---CGCTCCATCGAAGGCGATCTCAAACTCAATGGCAAGTCTCTCCGCAACAGCTTA 1631
Qy      1561  GAAGGCACTACCTCTAGCACTTTGGTCTCTGAGCAACACTGGGCTCTCCATCTCTGCA 1620
Db      1632  GAAGGCACTGCTCTAGCACTTTGGTCTCTGAGCAACACTGGGCTCTCCATCTCTGCA 1691
Qy      1621  GCCCGGCCAGCGCCCTCTAGGGGCTGTGAGTGAGGACAGCATCTGCTCCGAGTCC 1680
Db      1692  GCCCGGCCAGCGCTCTCTGGAGCTGTGAGTGAGGACAGCATCTGCTCCGAGTCC 1751
Qy      1681  TTTGACCAATTTGGAATCTGCTTGAACGCTCTTCCGAAACCCCACTGAGGGCTGTGTCT 1740
Db      1752  TTTGACCAATTTGGAATCTGCTTGAACGCTCTTCCGAAACCCCACTGAGGCTGTGTCT 1811
Qy      1741  GTGCAACACTGAGGGGCTTGAGCAGCTCTCCCTCAGAGTCTGAAGCGATGTGGCAG 1800
Db      1812  GTGCAACACTGAGGGGCTTGAGCAGCTCTCCCTCAGAGGCTTAAACGATGTGGCAG 1871
Qy      1801  GAATCTTTGGGGATAGTCTGCTTTCTGACAGCTGCAAGAGGTGACTGCGAGCTTAC 1860
Db      1872  GAATCTTTGGGGATAGTCTGCTTTCTGACAGCTGCAAGAGGTGACTGCGAGCTTAC 1931
Qy      1861  AGACAAGCCCTAGGAATCTGCTCAAGCTCAGC 1893
Db      1932  AGACAAGCCCTAGGAATCTGCTCAAGCTCAGC 1964

RESULT 6
AX380960
LOCUS      2929 bp      DNA      linear      PAT 18-MAR-2002
DEFINITION Sequence 1 from Patent WO0212456.
ACCESSION  AX380960
VERSION     AX380960.1  GI:19575800
KEYWORDS
SOURCE      Rattus sp.
ORGANISM    Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
AUTHORS    Drucker,D.J., Rosen,C.F. and Lefebvre,D.L.
TITLE      Ampk-related serine/threonine kinase, designated snark
JOURNAL    Patent: WO 0212456-A 1 14-FEB-2002;
            1149336 ONTARIO INC. (CA)
FEATURES   Location/Qualifiers
            source          1..2929
                        /organism="Rattus sp."
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:10118"

ORIGIN
Query Match      88.9%; Score 1683.4; DB 6; Length 2929;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 1769; Conservative 0; Mismatches 121; Indels 3; Gaps 1;

Qy      1  ATGAGTCCGTGGCTTACTCCAGCGCCGAGCAGGCTCCCTCGGCTCCGCGCTCGGCC 60
Db      83  ATGAGTCCGTGGCTTACTACCGCGCGGGAACCTGCTCCCTCGGCTCCGCGCTCGGCC 142
Qy      61  TCGGAGAGCGCCGCGCTGGGGAGCGGGCTCATCAAGTGGCTCAAAACCTCTGATGAAG 120

```

```

Db      143  ACGGAGAGCGCCCGCGCTGGCGGACCGGCTCATCAAGTCGCCCAACCTCTCATGAAG 202
Qy      121  AAGCAGCGGTGAAGCGGCAACCATCAAAACACAACCTGCGGCAACCGCTACGAGTCTCTG 180
Db      203  AAGCAGCGGTGAAGCGGCAACCATCAAAACACAACCTGAGGCAACCGCTACGAGTCTCTG 262
Qy      181  GAGACGCTGGGCAAGGCGCACTTACGGGAAGGTGAAGAGGCAAGAGAGCTCGGGCGGT 240
Db      263  GAGACGCTGGGCAAGGCGCACTTACGGGAAGGTGAAGAGGCAAGAGAGCTCGGGAGCG 322
Qy      241  CTGCTGGCCATCAAGTCCATCAGGAAAGACAAAATCAAGATGAGCAGGATCTCTGCAAC 300
Db      323  CTGCTGGCCATCAAGTCTATCAGGAGGACAAAATCAAGATGAGCAGGATCTCTGCAAC 382
Qy      301  ATACGAGGAGGATTTGAGATCATGTCTTCACTCAACCAACCCCAACATCATTTGCCATCAT 360
Db      383  ATAAGGAGGAGATTCGAGATCATGTCTTCACTCAACCAACCCCAACATCATTTGCCATCAT 442
Qy      361  GAAGTGTTCGAGATAGCAGCAAGATTTGATTTGTCATGAGTATGCCAGCGAGCGAT 420
Db      443  GAAGTGTTCGAGATAGCAGCAAGATTTGATTTGTCATGAGTATGCCAGCGAGCGAT 502
Qy      421  CTGTATGATTACATCAGTGAGCGGCGCACGCTGAGTGAGCGGAGCGGCGAGGCATTTCTTC 480
Db      503  CTGTACGATTACATCAGTGAGCGGCGCACGCTGAAATGAGCGGAGCGGCGAGGCATTTCTTC 562
Qy      481  CGACAGATCTGTCTGCGCTGCACTACTGTCACACAGAAACGGGATCTGTTTCAACGAGATCTC 540
Db      563  CGACAGATCTGTCTGCGCTGCACTACTGTCACACAGAAACGGGATTTGTTTCAACGAGATCTC 622
Qy      541  AAGCTGAAACATCTCTCTAGATGCCAATGGAACATCAAGATTCGTGAGTCTTGGGCTC 600
Db      623  AAGCTGAGAAATCTCTCTAGATGCCAATGGAACATCAAGATTCGTGAGTCTTGGGCTC 682
Qy      601  TCCAAACCTGTATCAAAAGGCAAGTTCTCTCAGACGTTCTGTGGGAGCCCTCTCTAGGCC 660
Db      683  TCCAAACCTGTATCAAAAGGCAAGTTCTCTCAGACGTTCTGTGGAGCCCTCTCTATGCC 742
Qy      661  TCGCTGAGATAGTCAACGGGAGCCCTATGTGGGCCAGAGGTGGACAGTGTGTCTG 720
Db      743  TCACCTGAGATCGTCAACGGGAGCCCTATGTGGGCCAGAGGTGGACAGTGTGTCTG 802
Qy      721  GCGTCTCTCTGACATCTCTGTCATGCGCACCATGCGCTTTGACGGGAGGATCATATA 780
Db      803  GCGTCTCTCTGACATCTCTGTCATGCGCACCATGCGCTTTGACGGGAGGATCATATA 862
Qy      781  ACACCTGTGAAGCAAAATCAGTAAACGGGCTTACCGTGAGCCGCCCAAGCCGTCGATGCC 840
Db      863  ACCCTGTGAAGCAAAATCAGTAAACGGGCTTACCGTGAGCCCGTGAACCCGTCGATGCC 922
Qy      841  TGTGGCTGATCCGGTGGCTGTTAATGTTGAACCCCAACCGCTCGGGCCACACCTGGAGAT 900
Db      923  TGTGGCTGATCCGGTGGCTGTTAATGTTGAATCCCATCCGTCGGGCGCACTCTGGAGAT 982
Qy      901  GTAGCCAGTCAATTTGGGTCGCTCAACCTGGGTTACACACCGGAGTCTGGGGAACAGGAAGCC 960
Db      983  GTAGCCAGTCAATTTGGGTCGCTCAACCTGGGTTACACACCGGAGTCTGGGGAACAGGAAGCT 1042
Qy      961  CTCGCTGAGGGTGGGCAACCTAGTGTGATCTTTGGCGGGCCCTCCATGGCGGAGCTGTTA 1020
Db      1043  CTCGAGAGGGTGGGCAACCTAGCGGTGATCTTGGCGGGCCCTCTATGGCGGAGCTGTTA 1102
Qy      1021  CGTGCCTCTCGCGCCCTCTCTGGAGATGGAGCCAGAGTGTGACGCTTCTTCAAGCAG 1080
Db      1103  CGTGCCTCTCTCGCGCCCTCTCTGGAGATGGAGCCAGAGTGTGACGCTTCTTCAAGCAG 1162
Qy      1081  CAGTGCCTGGGAGGTGAAGCACTGTACCTGGGCTGGAGCGGCAACATTTCTTAAAGAG 1140
Db      1163  CATGTGCGGGAGGTGAAGCACTGTACCTGGGCTGGAGCGGCTGGAGCGGCAACATTTCTTAAAGAG 1222
Qy      1141  TCCGGAAGGAGAGATGACATGGCTCAAAATCTGCAAGGTTGACCCGGCTGAGGATACCTCT 1200

```

```
Db 1223 TCCGCGAAGGAGANTGACATGGCTCAGACTCTGCAGATGACCCAGTTGAAGATCTTCC 1282
Qy 1201 TCTCGCCCTGGCAAGAGCAGCTTTAAGCTTCCGAAAGGCATTTCTCAAGAAAAGTCTCTCT 1260
Db 1283 TCTCGCCCTGGCAAGAGCAGCTCAAGCTTCCGAAAGGTATCTCTCAAGAAAAGGCTCTCT 1342
Qy 1261 ACCTGCTCAGGGAGGTACAGGAGGACCTCTCAGGAATCTCAGACCGGTGCTGATCTCCA 1320
Db 1343 CCCTCATCGGGGAGGTACAGGAGGCGCTCTCAGGAATCTCAGACCGAGTGTCCAATACCCCA 1402
Qy 1321 GGGCAGCTGTCCTGCTGTATCTCTGCTCCCAAGGAAAGGCATCTTAAAGACTCTCGA 1380
Db 1403 GGGCAGCTGTCCTGCTGTATCTCTGCTCCCAAGGAAAGGCATCTTAAAGACTCTCGG 1462
Qy 1381 CAGCGTGAATCTGTTACTACTCTCTCCAGAGCCCGAGTCTGGGGAATCTTAGAC 1440
Db 1463 CAGCGTGAATCTGTTACTACTCTCTCCAGAGCCCGAGTCTGGGGAATCTTAGAC 1522
Qy 1441 GCCAGTGAATGTTGTTGAGTGGGAGCCCGTGGAGCAGAAAGTCTCCACAGGCTTCAGGG 1500
Db 1523 GCAGGTGATGTTGTTGAGTGGGAGCCCGTGGAGCAGAAAGTCTCCACAGGCTTCAGGG 1582
Qy 1501 CTCCTCTCCACCGCAGGAGTCTCAAACTCAATGCGAAGTCTCCCGCAGGCTTCA 1560
Db 1583 ---CGCCTCCATCGCAAGGAGCATCTCAAACTCAATGCGAAGTCTCCCGCAGGCTTCA 1639
Qy 1561 GAAGGCACTACCCCTAGCACCTTTGGCTCTCCGACCAACTGGGCTCTCCCATCTCTGCA 1620
Db 1640 GAAGGCACTGCCCTAGCACCTTTGGCTCTCCGACCAACTGGGCTCTCCCATCTCTGCA 1699
Qy 1621 GCCGGGCCAGCCGCCCTCAGGGGCTGTGAGTGAGCAGCAGCATCTGTCTCCGAGTCC 1680
Db 1700 GCCGGGCCAGCCGCCCTCAGGGAGCTGTGAGTGAGCAGCAGCATCTGTCTCCGAGTCC 1759
Qy 1681 TTTGACCAATTTGACTTGGCTGACGCTTCCGACCAACTGGGCTCTCCCATCTCTGCTCT 1740
Db 1760 TTTGACCAATTTGACTTGGCTGACGCTTCCGACCAACTGGGCTCTCCCATCTCTGCTCT 1819
Qy 1741 GTGACCAACTGAGGGGCTTGGACGAGCTCTCAGAGTCTGAGGAGCATCTGTCTCCGAGTCC 1800
Db 1820 GTGACCAACTGAGGGGCTTGGACGAGCTCTCAGAGTCTGAGGAGCATCTGTCTCCGAGTCC 1879
Qy 1801 GAATCTTTGGGGATAGTGTCTTTCTGACAGACTGCCAAGAGGTGACTGCGAGCCTAC 1860
Db 1880 GAATCTTTGGGGATAGTGTCTTTCTGACAGACTGCCAAGAGGTGACTGCGAGCCTAC 1939
Qy 1861 AGACAAGCCCTAGGAATCTCTCAAGCTCAGC 1893
Db 1940 AGACAAGCCCTAGGAATCTCTCGAAGCTCAGC 1972

RESULT 7
AC125887 LOCUS 291762 bp DNA linear HTG 15-NOV-2002
DEFINITION Rattus norvegicus clone CH230-274J3, *** SEQUENCING IN PROGRESS
***, 13 unordered pieces.
AC125887
VERSION AC125887.3 GI:25008671
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Rattus.
1 (bases 1 to 291762)
Muzny,D,Marle, Metzker,M, Lee, Abramzon, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
```

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Direct Submission  
Unpublished  
2 (bases 1 to 291762)  
Worley, K.C.  
Direct Submission  
Submitted (02-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 291762)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Nov 15, 2002 this sequence version replaced gi:23269681.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

```
----- Project Information
Center project name: GWLD
Center clone name: CH230-274J3
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 226053 bases at least Q40
Consensus quality: 230853 bases at least Q30
Consensus quality: 233437 bases at least Q20
Estimated insert size: 233445; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 13 contigs. The true order of the pieces
  is not known and their order in this sequence record is
  arbitrary. Gaps between the contigs are represented as
  runs of N, but the exact sizes of the gaps are unknown.
  This record will be updated with the finished sequence
  as soon as it is available and the accession number will
  be preserved.
* 1 70277: contig of 70277 bp in length
* 70278 70377: gap of unknown length
* 95031: contig of 24654 bp in length
* 95032 95131: gap of unknown length
* 108956: contig of 13825 bp in length
* 108957 109056: gap of unknown length
* 109057 271697: contig of 162641 bp in length
* 271698 271797: gap of unknown length
* 271798 273026: contig of 1229 bp in length
* 273027 273126: gap of unknown length
* 273127 274531: contig of 1405 bp in length
* 274532 274631: gap of unknown length
* 274632 276287: contig of 1656 bp in length
* 276288 276387: gap of unknown length
* 276388 277904: contig of 1517 bp in length
* 277905 278004: gap of unknown length
* 278005 279523: contig of 1519 bp in length
* 279524 279623: gap of unknown length
* 279624 280724: contig of 1101 bp in length
* 280725 280824: gap of unknown length
* 280825 284658: contig of 3834 bp in length
* 284659 284758: gap of unknown length
* 284759 287234: contig of 2476 bp in length
* 287235 287334: gap of unknown length
* 287335 291762: contig of 4428 bp in length.
----- Location/Qualifiers
  1. 291762
     /organism="Rattus norvegicus"
     /mol_type="genomic DNA"
     /db_xref="taxon:10116"
     /clone="CH230-274J3"
     70378..75596
        /note="wgs_contig"
     95132..96554
        /note="wgs_contig"
     109057..110574
        /note="wgs_contig"
-----
Query Match      80.7%; Score 1528.4; DB 2; Length 291762;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1535; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
-----
QY 1 ATGAGTCGTGGCTTACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCGGCTCGGCC 60
Db 274742 ATGAGTCGTGGCTTACTCCAGCGCCGAGCCAGGCTCCCTCGGCTCGGCTCGGCC 274801
QY 61 TCGGAGAGCGCCGCGCTGGCGGACGGGCTCATCAAGTCGCCTAAACCTCTGATGAAG 120
Db 274802 TCGGAGAGCGCCGCGCTGGCGGACGGGCTCATCAAGTCGCCTAAACCTCTGATGAAG 274861
QY 121 AAGCAGCGGTGAAGCGGCACCATCAAAACAACTCGGCGCACCGCTACGAGTTCCTG 180
```

```
Db 274862 AAGCAGCGGTGAAGCGGCACCATCAAAACAACTCGGCGCACCGCTACGAGTTCCTG 274921
QY 181 GAGACGCTGGGCAAGGCACTTACGGGAAGGTGAAGAACGACGAGAGCTCGGGGCGT 240
Db 274922 GAGACCTGGGCAAGGCACTTACGGGAAGGTGAAGAACGACGAGAGCTCGGGGCGT 274981
QY 241 CTGGTGGCCATCAAGTCCATCAGGAAGACAAAATCAAGATGAGCAGGATCTGTGTCAC 300
Db 274982 CTGGTGGCCATCAAGTCAATCAGGAAGACAAAATCAAGATGAGCAGGATCTGTGTCAC 275041
QY 301 ATACGAGGAGATGAGATCATGTCTTCACTCAACACCCCAATCATTCATTCGATCCAT 360
Db 275042 ATACGAGGAGATGAGATCATGTCTTCACTCAACACCCCAATCATTCGATCCAT 275101
QY 361 GAAAGTGTTCGAGAAATAGCAAGATTTGATTTGTCTATGAGATGATGCACCGAGGCGAT 420
Db 275102 GAAAGTGTTCGAGAAATAGCAAGATTTGATTTGTCTATGAGATGATGCACCGAGGCGAT 275161
QY 421 CTGTATGATTAATCATAGTAGCGGCGCACGGCTGAGTGAGCGGACGCGCAGGCATTTCTTC 480
Db 275162 CTGTATGATTAATCATAGTAGCGGCGCACGGCTGAGTGAGCGGACGCGCAGGCATTTCTTC 275221
QY 481 CGACAGATCGTGTCTGCTCTGCACTACTGCCACCAAGAGCGGATCGTTCAACCGAGATCTC 540
Db 275222 CGACAGATCGTGTCTGCTCTGCACTACTGCCACCAAGAGCGGATCGTTCAACCGAGATCTC 275281
QY 541 AAGCTGGAAACATCTCTTCTAGATGCAATGGAACATCAAGATTTGCTGATTTGGCCTC 600
Db 275282 AAGCTGGAAACATCTCTTCTAGATGCAATGGAACATCAAGATTTGCTGATTTGGCCTC 275341
QY 601 TCCAACTGTACCAAGCAAGTTCTTCCAGACGTTCTGTGGGAGCCCTCTCTAGGCC 660
Db 275342 TCCAACTGTACCAAGCAAGTTCTTCCAGACGTTCTGTGGGAGCCCTCTCTAGGCC 275401
QY 661 TCCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCAGAGAGTGACAGCTGCTCTCTG 720
Db 275402 TCCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCAGAGAGTGACAGCTGCTCTCTG 275461
QY 721 GCGGTCTCTCTGATCATTCCTGTGTCATGGCAATGCGCTTTTCAACGGGAGGATCATATA 780
Db 275462 GCGGTCTCTCTGATCATTCCTGTGTCATGGCAATGCGCTTTTCAACGGGAGGATCATATA 275521
QY 781 AACTGTGTAAGCAATCAAGTACGGGGCTTACCGTGAGCGCCCGCAGCGCTCGGATGCC 840
Db 275522 AACTGTGTAAGCAATCAAGTACGGGGCTTACCGTGAGCGCCCGCAGCGCTCGGATGCC 275581
QY 841 TGTGGCTGATCCGCTGGCTGTTAAATGGTGAAACCCCAACCCCGTCCGGCCACACCTGGAGGAT 900
Db 275582 TGTGGCTGATCCGCTGGCTGTTAAATGGTGAAACCCCAACCCCGTCCGGCCACACCTGGAGGAT 275641
QY 901 GTAGCAGTCAATTTGGTGGGTCAACTGGGGTTACACCAACCGAGTGGGGAACAGGAAGCC 960
Db 275642 GTAGCAGTCAATTTGGTGGGTCAACTGGGGTTACACCAACCGAGTGGGGAACAGGAAGCC 275701
QY 961 CTGCGTGAGGTTGGGCAACCTTAGTGTGACTTTTGGCGGCGCCCTCCATGGCGGACTGTTA 1020
Db 275702 CTGCGTGAGGTTGGGCAACCTTAGTGTGACTTTTGGCGGCGCCCTCCATGGCGGACTGTTA 275761
QY 1021 CFTCGCTCTCGCGCCCTCTCTTGAGAAATGGAGCAAGGTGTGACGCTTCTTCAAGCAG 1080
Db 275762 CFTCGCTCTCGCGCCCTCTCTTGAGAAATGGAGCAAGGTGTGACGCTTCTTCAAGCAG 275821
QY 1081 CAGCTCCGGGAGGTGAAGCACTGTACTGGCTGAGCGGCAACATTTCTTTAAGAAG 1140
Db 275822 CAGCTCCGGGAGGTGAAGCAACCGTACCTGGCTGAGCGGCAACATTTCTTTAAGAAG 275881
QY 1141 TCCGAAAGGAGAAATGACATGGCTCAAAATCTGAAGGTGACCCGGCTGAGGATACCTCT 1200
Db 275882 TCCGAAAGGAGAAATGACATGGCTCAAGATCTGAAGGTGACCCGGCTGAGGATACCTCT 275941
QY 1201 TCTCGCCCTGGCAAGAGCGCTTAAAGCTTCGAAAGGCAATTTCTCAAGAAAAAGTCTCTCT 1260
```

Db 275942 TCTCGTCTGCGAGAGAGCGCTTAAGCTTCCAAAAGGCATTTCTCAAGAAAAGTCTCTCT 276001  
Qy 1261 ACCTCGTCAGGGGAGGTACAGAGGAGCCCTCAGGAACTCAGACGGTGCTCTGATCTCCA 1320  
Db 276002 ACCTCTTCAGGGGAGGTACAGAGGAGCCCTCAGGAACTCAGACGGTGCTCTGATCTCCA 276061  
Qy 1321 GGGCAGCTGTCCCTGCTGTATCCCTGCTCCAGGAAAGGCATCTTAAGAGTCTCGA 1380  
Db 276062 GGGCAGCTGTCCCTGCTGTATCCCTGCTCCAGGAAAGGCATCTTAAGAGTCTCGA 276121  
Qy 1381 CAGCGTGAATCTGGTACTACTCTCTCCAGAGCCAGCGAGTCTGGGGAAGTCTTAGAC 1440  
Db 276122 CAGCGTGAATCTGGTACTACTCTCTCCAGAGCCAGCGAGTCTGGGGAAGTCTTAGAC 276181  
Qy 1441 GCCAGTGAATGTTGTGAGTGGGAGCCCGTGAGGAGAGTCTTCCACAGGCTTCAGGG 1500  
Db 276182 GCCAGTGAATGTTGTGAGTGGGAGCCCGTGAGGAGAGTCTTCCACAGGCTTCAGGG 276241  
Qy 1501 CTCTCTCTCCAGGAGGAGTCTCAAACTCAATGGCAAGTCT 1546  
Db 276242 CTCTCTCTCCAGGAGGAGTCTCAAACTCAATGGCAAGTCT 276287

RESULT 8  
AX407118

LOCUS AX407118 1884 bp DNA linear PAT 14-JUN-2002  
DEFINITION Sequence 3 from Patent WO0224921.  
ACCESSION AX407118  
VERSION AX407118.1 GI:21439883  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Curtis, R.A. and Galvin, K.M.  
3700, a novel human protein kinase and uses therefor  
Patent: WO 0224921-A 3 28-MAR-2002;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
Location/Qualifiers  
1..1884  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 71.9%; Score 1360.2; DB 6; Length 1884;  
Best Local Similarity 84.2%; Pred. No. 0;  
Matches 1565; Conservative 0; Mismatches 273; Indels 21; Gaps 2;

Qy 47 CCTCCGCGCTCGAGAGCGCCCGCGCTGGCGGAGCGGCTCATCAAGTCGCTA 106  
Db 35 CCACTCCCTCGCGCGAGAGCTAGCCGCGCTGGCGGAGGCTGATCAAGTCGCCA 94  
Qy 107 AACCTCTGATGAAGACAGCGCGTGAAGCGGACCATCAACAACAACCTCGCGCAC 166  
Db 95 AGCCCTTAATGAAGACAGCGCGTGAAGCGGACCAACCAAGCAAACTCGCGCAC 154  
Qy 167 GCTACAGATTCCTGGAGACCTGGGCAAGGCGACCTACGGGAGGTGAAGAGCCAG 226  
Db 155 GCTACAGATTCCTGGAGACCTGGGCAAGGCGACCTACGGGAGGTGAAGAGCGCGG 214  
Qy 227 AGAGCTCGGCGCTCTGGTGGCATCAAGTCCATCAGGAAGACAAATCAAGATGAGC 286  
Db 215 AGAGCTCGGCGCTCTGGTGGCATCAAGTCCATCAGGAAGGACAAATCAAGATGAGC 274  
Qy 287 AGATCTGTGCATACAGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCA 346  
Db 275 AAGATCTGTGCATACAGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCA 334  
Qy 347 TCATTGCCATCCATGAAGTGTGAGATGAGGAGAGATTGATGATGATGATGATGATG 406  
Db 335 TCATTGCCATCCATGAAGTGTGAGATGAGGAGAGATTGATGATGATGATGATGATG 394

Qy 407 CCAGCCGAGGCGATCTGTATGATTACATCAGTGAGCGGCCACGGCTGAGTCAGCGGGACG 466  
Db 395 CCAGCCGAGGCGATCTGTATGATTACATCAGTGAGCGGCCACGGCTGAGTCAGCGGGACG 454  
Qy 467 CCAGGCAATTTCTCCGACAGATCTGTGTCGCCCTGCACTACTGCAACAGAAACGGGATCG 526  
Db 455 CTAGGCAATTTCTCCGACAGATCTGTGTCGCCCTGCACTACTGCAACAGAAACGGGATCG 514  
Qy 527 TTCACCCAGATCTCAAGCTGGAAACATCTCTCTAGATGCCAATGGAAACATCAAGATTG 586  
Db 515 TCCACCCAGATCTCAAGCTGGAAACATCTCTCTGGATGCCAATGGAAATATCAAGATTG 574  
Qy 587 CTGACTTTGGCCTCTCCAAACCTGTACCAAGCAAGTTCTCTCAGACGTTCTGTGGGA 646  
Db 575 CTGACTTTGGCCTCTCCAAACCTGTACCAAGCAAGTTCTCTCAGACATCTCTGTGGGA 634  
Qy 647 GCCCTCTCTACGCTCGCCTGAGATGATCAAGGGAAAGCCCTATGTGGGCCAGAGATGG 706  
Db 635 GCCCTCTCTATGCTCGCCAGAGATTGTCAATGGAAAGCCCTACACAGGCCAGAGGTGG 694  
Qy 707 ACAGCTGTCTCTGGGCGTTCTCTGTACATCTCTGTGCAATGGCAACATGCGCTTTGACG 766  
Db 695 ACAGCTGTCTCTGGGCGTTCTCTGTACATCTCTGTGCAATGGCAACATGCGCTTTGATG 754  
Qy 767 GGCAGGATCATAAACACACTGCTGAAGCAAAATCAGTAACGGGGCTTACCGTGAGCGGCCA 826  
Db 755 GGCATGACCATATAGATCTTAGTGAACAGATCAGCAACGGGGCTTACCGGAGGCCACTA 814  
Qy 827 AGCGTCCGATGCTGTGGCTGATCCGCTGATCCGCTGATTAATGGTGAACCCCAACCGTCGG 886  
Db 815 AACCTCTGATGCTGTGGCTGATCCGCTGATCCGCTGATTAATGGTGAACCCCAACCGTCGG 874  
Qy 887 CCACACTGGAGGATGATGAGCAGATCTGTTGGGTCAACTGGGGTTACACCAACCGAGTCG 946  
Db 875 CCACCTTGGAGGATGTTGGCAGTCACTGTTGGGTCAACTGGGGTTACCGCCACCGAGTGG 934  
Qy 947 GGGAAACAGGAAGCCCTCGCTGAGGGTGGGCAACCTAGTGTGACTTTTGGCGGGGCCCTCCA 1006  
Db 935 GAGAGCAGGAGGCTCCGCATGAGGGTGGGCAACCTGCGAGTACTCTGCCCGGCCCTCCA 994  
Qy 1007 TGGCGGACTGTTACGTCGCTCTCGCGCCCTCTCTGGAGAAATGGAGCCAAAGTGTGCA 1066  
Db 995 TGGCTGACTGGCTCGGCGTTCTCCCGCCCTCTCTGGAGAAATGGGCGCAAGGTGTGCA 1054  
Qy 1067 GCTTCTCAAGCAGCAGCTGCGCGGAGGTGGAGACACTGTACCTGGGCTGGAGCGGCAC 1126  
Db 1055 GCTTCTTCAAGCAGCAGTGCACCTGGTGGGGAAGCACACCCCTGGGCTGGAGCGCCAGC 1114  
Qy 1127 ATTCTCTTAAGAAAGTCCCGAAAGAGAAATGACATGGCTCAAAAATCTCAAGGTGACCCGG 1186  
Db 1115 ATTGCTCAAGAAAGTCCCGCAAGGAGAAATGACATGGCCAGTCTCTCCACAGTGACAGG 1174  
Qy 1187 CTGAGGATACCTTTCTCGCCCTGGCAAGAGCAGCTTAAGCTTCCGAAAGGCATTTCTCA 1246  
Db 1175 CTGATGACACTGCGCATCGCCCTGGCAAGAGCAACCTCAAGCTGCCAAAGGCAATTTCTCA 1234  
Qy 1247 AGAAAGTCTCTACTCTCTGTCAGGGGAGGTACAGGAGGACCTCTCAGGAATCTCAGACGG 1306  
Db 1235 AGAAAGTCTCTCAGCTCTTGCAGAAAGGGGTACAGGAGGACCTCTCCGAGCTCAGCCCAA 1294  
Qy 1307 TGCTGATACTCCAGGCGACCTGTCTCTGCTGTATCCCTGCTCCCTCCCAAGGAAGGCATCC 1366  
Db 1295 TCCCTGGAGCCAGGCGAGGCTGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1345  
Qy 1367 TTAAGAGTCTCAGACAGCGTGAATCTGTGTTACTACTCTCTCTCAGAGCCAGCGAGTCTG 1426  
Db 1346 TCAAGAGCCCGCAGACAGCGAGTCTGGGTACTACTCTCTCTCTCTCTCTCTCTCTCTCTCT 1405  
Qy 1427 GGGAACTCTTAGAGCCAGTGTGTTGTGAGTGGGAGCCCGCTGGAGGAGAGTCTC 1486  
Db 1406 GGGAGCTCTTGGAGCGAGCGAGCTGTTTGTGAGTGGGAGTCCCAAGGAGCAGAAAGCTC 1465

```
QY 1487 CACAGGCTTCAGGCTCCTCTCCACCGCAAGGCAATTCCTCAAACTCAATGGCAAGTTCT 1546
DB 1466 CGAAGCTTCAGGCTCCTCTCCATCGAAAGGATCCTCAAACTCAATGGCAAGTTCT 1525
QY 1547 CCGCACAGCCTTAGAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCCT 1606
DB 1526 CCACAGCAGCCTTAGAGCTCGGGGCCCCACCACTTCCTGGCTCCCTGGATGAATCGCCC 1585
QY 1607 CTTCCCATCTGCAGCCCGGCCCGAGCCCTCAGGGCTGTGAGTGAGGACAGATCC 1666
DB 1586 CACCTCGCCCTCGGCCCGGCCAGCCCTCAGGGCTGTGAGGAGGACAGATCC 1645
QY 1667 TGTCTCCGAGTCTTTGACCAATTTGACTTTGCTGAAGCTCTTCCGAAACCCACTGA 1726
DB 1646 TGTCTCTGAGTCTTTGACCACTGAGTCTTGTGCTGAACTGCTCCAGAGCCCTCCACTGC 1705
QY 1727 GGGCTGTGTCTGTGGCAAACTGAGGGGCTTGGAGCCTCCCTCGAAG- 1780
DB 1706 GGGCTGTGTCTGTGGCAAACTCAGGGGCTTGGAGGCCCCCTCAGAGGGCCCTG 1765
QY 1781 -----GTCTGAAGCGATGTGGCAGGAATCTTTGGGGATAGCTGCTTTCTGTGACAG 1834
DB 1766 GAAGTGTCTGTGAGCGCTGCGGCGAGGATCTTTGGGGGACAGCTGCTTTTCCCTGACAG 1825
QY 1835 ACTGCCAAGAGGTGACTGACGCTTACAGACAGCCCTAGGAATCTGCTCAAGCTCAGC 1893
DB 1826 ACTGCCAAGAGGTGACAGCACTTACAGAGGCACTGAGGCTGTCTCAAGCTCACC 1884

RESULT 9
LOCUS AX407116 3353 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 1 from Patent WO0224921.
ACCESSION AX407116
VERSION AX407116.1 GI:21439882
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Curtis, R.A. and Galvin, K.M.
3700, a novel human protein kinase and uses therefor
Patent: WO 0224921-A 1 28-MAR-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Source
1. 3353
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 71.9%; Score 1360.2; DB 6; Length 3353;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 273; Indels 21; Gaps 2;

QY 47 CCTCCGCTCGGCTCGAGAGCCCGCGCTGGCGGAGCGGCTCATCAAGTCGCTA 106
DB 191 CCATCCTCGGCGCAGAGTAGCCCGCGCTGGCGGAGGGCTGATCAAGTCGCCCCA 250
QY 107 AACCTCTGATGAAGAAGCAGCGGTGAAGCGGCACCATCAAAACACAACCTCGGCGACC 166
DB 251 AGCCCTTAATGAAGAAGCAGCGGTGAAGCGGCACCAACCAAGCAACCTCGGCGACC 310
QY 167 GTTACGAGTTCTGGAGACGCTGGGCAAGGCACTTACGGGAAGGTGAAGGACAGAG 226
DB 311 GCTACGAGTTCTGGAGACCTGGGCAAGGCACTTACGGGAAGGTGAAGGAGCGGG 370
QY 227 AGAGTCGGGCGTCTGTGTCATCAAGTCCATCAGGAAGACAAATCAAGATGAGC 286
DB 371 AGAGTCGGGCGCTGTGTGTCATCAAGTCCATCAGGAAGACAAATCAAGATGAGC 430
QY 287 AGGATCTGTGCATACGAGGAGGATGAGATCATGTCTTCACTCAACACCCCCACA 346
```

```
DB 431 AAGATCTGATGCACATACGAGGAGATGAGATCATGTCTCATCAACCAACCTCACA 490
QY 347 TCATTGCTCATGATGATGTTTGAATAGCAGCAAGATGATGATGATGATGATGATGATG 406
DB 491 TCATTGCTCATGATGATGTTTGAATAGCAGCAAGATGATGATGATGATGATGATGATG 550
QY 407 CCAGCGAGGAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 466
DB 551 CCAGCGGAGGAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 610
QY 467 CCAGCGATTTCTTCGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 526
DB 611 CTAGGATTTCTTCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 670
QY 527 TTACCGAGATCTCAAGCTGGAATAATCTCTTAGATGATGATGATGATGATGATGATGATG 586
DB 671 TCACCGAGATCTCAAGCTGGAATAATCTCTTAGATGATGATGATGATGATGATGATGATG 730
QY 587 CTGACTTTGGCTCTCCAACTGTACCAAAAGCAAGTTCTTCCAGACGTTCTGTGGGA 646
DB 731 CTGACTTTGGCTCTCCAACTGTACCAAAAGCAAGTTCTTCCAGACGTTCTGTGGGA 790
QY 647 GCCTCTCTACGGCTCGCTGAGATGATCAACGGAAGCCCTATGTGGGCCCAAGAGTGG 706
DB 791 GCCCTCTCTATGCTCGCAGAGATGATCAATGGGAAGCCCTACACAGGCCCAAGAGTGG 850
QY 707 ACAGCTGCTCTGTGGGCTTCTCTGTATATCTCTGTGATGATGATGATGATGATGATGATG 766
DB 851 ACAGCTGCTCTGTGGGCTTCTCTGTATATCTCTGTGATGATGATGATGATGATGATGATG 910
QY 767 GGCAAGATCATAAACACACTGTGTAAGCAAACTAGTAACGGGCTTACCGTGAGCGGCCA 826
DB 911 GGCAAGATCATAAACACACTGTGTAAGCAAACTAGTAACGGGCTTACCGTGAGCGGCCA 970
QY 827 AGCGCTCCGATGCTGTGGCTGATCCGCTGGCTGATCCGCTGGCTGATCCGCTGGCTGAT 886
DB 971 AACCTCTGTGCTGTGGCTGATCCGCTGGCTGATCCGCTGGCTGATCCGCTGGCTGAT 1030
QY 887 CCACACTGGAGATGATGAGCAGCTATGTTGGTGGCTCACTGGGTTTACCAACCGGAGTCG 946
DB 1031 CCACCTTGGAGATGATGAGCAGCTATGTTGGTGGCTCACTGGGTTTACCAACCGGAGTCG 1090
QY 947 GGGAAAGAGGAGCTCGCTGAGGCTGGGCAACCTAGTGTGATCTTTGGCGGGCTCCCA 1006
DB 1091 GAGAGCAGGAGGCTCGCAGATGAGGCTGGGCAACCTGTCAGTGTGATCTGCGCGCTCCA 1150
QY 1007 TGGCGGACTGTTTACGTCGCTCTCGCGCCCTCTCTGGAGAAATGAGGACCAAGTGTGCA 1066
DB 1151 TGGCTGACTGGCTCGGCGCTTCTCTCGCGCCCTCTCTGGAGAAATGAGGACCAAGTGTGCA 1210
QY 1067 GCTTCTTCAAGCAGCAGCTGCGGAGGTGGAAGCACTGTACCTGGCTGGAGCGGCAAC 1126
DB 1211 GCTTCTTCAAGCAGCAGCTGCGGAGGTGGAAGCACTGTACCTGGCTGGAGCGGCAAC 1270
QY 1127 ATTCTCTTAAGAAAGTCCGAAAGGAGAAATGACATGGCTCAAAATCTCAAGAGTCAACCGG 1186
DB 1271 ATTCTCTTAAGAAAGTCCGAAAGGAGAAATGACATGGCTCAAAATCTCAAGAGTCAACCGG 1330
QY 1187 CTGAGGATACCTCTTCTGCGCTTGGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATCTCA 1246
DB 1331 CTGATGACACTGCGCTTGGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATCTCA 1390
QY 1247 AGAAAGAGTCTCTACCTCTCAGGGAGGTACAGGAGGACCTTCAGGAATCTCAGACCGG 1306
DB 1391 AGAAAGAGTCTCTACCTCTCAGGGAGGTACAGGAGGACCTTCAGGAATCTCAGACCGG 1450
QY 1307 TGCCTGATCTCAGGCGCAGCTGTCTCTGTGTATCCCTGCTCTCCCAAGGAAAGGATCC 1366
DB 1451 TCCCTGCGAGCCCGAGGCGAGGCTGCC- 1501
QY 1367 TTAAGAAAGTCTCAGCAGCGTGAATCTGGTTACTTCTCTCTCAGAGCCCGAGCTGTG 1426
```

Db 1502 TCAAGAAAGCCCGCAGCGAGTCTGGCTACTACTCTCTCCGAGCCCAAGTGAATCTG 1561  
Qy 1427 GGGAACTCTTAGACGCGAGTATGTTTGTAGTGGGACCCCGTGGAGCAGAGTCTC 1486  
Db 1562 GGGAGCTCTTGGAGCGAGCGAGCTGTTTGTAGTGGGATCCCAAGAGCAGAGCCCTC 1621  
Qy 1487 CACAGGCTTCAGGCTCTCTCTCCACCGCAAGGCAATCTCAAACTCAATGGCAAGTTCT 1546  
Db 1622 CGCAAGCTTCAGGCTCTCTCTCCATCGCAAGGCAATCTCAAACTCAATGGCAAGTTCT 1681  
Qy 1547 CCGCAGAGCTTAGAGGCACTACCTCTAGCAGCTTTGGCTCCCTGGACCAACTGGCCT 1606  
Db 1682 CCCAGACAGCTTAGAGCTCGCGGCCCCACACCTTCGGCTCCCTGGATGAATCGGCC 1741  
Qy 1607 CTTCCCTCTGAGCGCGCCGCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCC 1666  
Db 1742 CACTCGCCCTCGCGCGGCCGAGCCCTCAGGGGCTGTGAGCGAGGACAGCATCC 1801  
Qy 1667 TGTCTCCGAGCTCTTTGACCAATTTGGACTTTGCCTGAACGTTCTTCCGAAACCCCACTGA 1726  
Db 1802 TGTCTCTGAGTCTTTGACCACTGAGCTTTGCCTGAACGCTCCAGAGCCCCCACTGC 1861  
Qy 1727 GGGCTGTGTCTGTGTGACAACTGAGGGGCTTTGAGCAGCCCTCCCTCAGAAG----- 1780  
Db 1862 GGGCTGTGTCTGTGTGACAACTCAGGGGCTTTGAGGAGCCCTCAGAGGGGCCCTG 1921  
Qy 1781 -----GTCGAAGCGATGTGGGAGGATCTTTGGGGATAGTCTTTCTGTGACAG 1834  
Db 1922 GAAGCTGCTGAGGCGCTGGCGGAGGATCTTTGGGGGAGAGTGTCTTTCCCTGACAG 1981  
Qy 1835 ACTGCCAAGAGTGAAGTGCAGCTTACAGACGCTTAGGAATCTGCTCAAGAGCTCAGC 1893  
Db 1982 ACTGCCAAGAGTGAAGTGCAGCTTACAGACGCTTAGGAATCTGCTCAAGAGTCAAC 2040

## RESULT 10

AX642961 AX642961 3360 bp DNA linear PAT 24-FEB-2003  
LOCUS Sequence 38 from Patent WO01096547.  
ACCESSION AX642961

VERSION AX642961.1 GI:28550104

## KEYWORDS

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1

## AUTHORS

Yue, H., Lal, P., Bandman, O., Borowsky, M., Au-Young, J., Lu, Y.,  
Gandhi, A. R., Tribouley, C. M., Wallia, N., Yao, M. G., Lu, D. A.,  
Greenwald, S. R., Rankumar, J., Griffin, J. A., Kearney, L., Burford, N.,  
Nguyen, D. B., Tang, Y. T., Baughn, M. R., He, A., Thornton, M.,  
Hafalia, A., Patterson, C., Gururajan, R., Lo, T. P., Khan, F.,  
Recipon, S. A., Asimail, Y., Policky, J. L., Ding, L., Grether, M.,  
Elliott, V. S., Thangavelu, K., Batra, S. and Ison, C. H.

## TITLE

Humain kinases

## JOURNAL

Patent: WO 01096547-A 38 20-DEC-2001;

## FEATURES

## source

1. 3360  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 4841542CB1"

## ORIGIN

Query Match 71.7%; Score 1357.6; DB 6; Length 3360;  
Best Local Similarity 83.7%; Pred. No. 0;  
Matches 1569; Conservative 0; Mismatches 284; Indels 21; Gaps 2;  
Qy 32 GCCAGGCTCCCTCGGCTTCGCTTGGCTCGGAGAGCGCCCGCTGGCGAGCGGC 91  
Db 98 CGCGGGCGCTCCGCGCCCACTCCCTCGCGGAGAGTAGCCCGCGCTGGCGAAGGC 157

Qy 92 TCATCAAGTGCCTAAACCTCTGATGAAGAGCAGCGGTGAAGCGGCACCATCACAAAC 151  
Db 158 TGATCAAGTGCCTAAACCTCTGATGAAGAGCAGCGGTGAAGCGGCACCATCACAAAC 217  
Qy 152 ACAACCTGCAGCACCCTACGAGTTCCTGGAGACGCTGGGCAAGGGCACTTACGGGAAG 211  
Db 218 ACAACCTGCAGCACCCTACGAGTTCCTGGAGACGCTGGGCAAGGGCACTTACGGGAAG 277  
Qy 212 TGAGAGGACACGAGAGGCTCGGGGCTGTGGTGGCCATCAAGTCCATCAGGAAGACA 271  
Db 278 TGAAGAGGCGCGGAGAGCTCGGGGCGCCCTGGTGGCCATCAAGTCAATTCGGAAGGACA 337  
Qy 272 AAATCAAGATGAGCAGGATCTGCTGCACATACCGAGGGAGATTGAGATCATGTCTTCAC 331  
Db 338 AAATCAAGATGAGCAGGATCTGATGCACATACCGAGGGAGATTGAGATCATGTTCATCG 397  
Qy 332 TCAACCAACCCCAATCATTTGCCATCCATGAAGTGTGAGAAATAGCAGCAAGATTGTGA 391  
Db 398 TCAACCAACCCCTCAATCATTTGCCATCCATGAAGTGTGAGAAACAGCAGCAAGATCGTGA 457  
Qy 392 TTGTCATGGAGTATGCCAGCGGAGCGATCTGTATGATTATCATCAGTGAGCGGCCACGGC 451  
Db 458 TCGTCATGGAGTATGCCAGCGGCGGCGACCTTTATGACTACATCAGCGAGCGGCGAGCAG 517  
Qy 452 TGAGTGAGCGGAGCGCGCAGGCAATTTCTTCCGACAGATCGTGTCTGCCCTGCACCTACTGCC 511  
Db 518 TCAGTGAGCGGAGCGCTAGGCAATTTCTTCCGACAGATCGTGTCTGCCCTGCACCTACTGCC 577  
Qy 512 ACCAAGACGGGATCGTTTCAACGAGATCTCAAGCTGGAAACATCTCTTCTAGATGCCAATG 571  
Db 578 ATCAGAACAGAGTTGTCCACCGAGATCTCAAGCTGGAGAACATCTCTTCTGGTGGCAATG 637  
Qy 572 GAAACATCAAGATTGCTGACTTTGGCCCTCTCCAACTGTACCAAAAGGCAAGTTCCCTCC 631  
Db 638 GGAATATCAAGATTGCTGACTTTGGCCCTCTCCAACTGTACCAAAAGGCAAGTTCCCTGC 697  
Qy 632 AGAGCTGTGGGAGCGCTCTCTACGCTCGCTGAGATAGTCAAGGGAGCGCTTATG 691  
Db 698 AGACATTTCTGGGAGCGCCCTCTATGCTTCCGACAGAGATTGTCAATGGGAGCGCTTACA 757  
Qy 692 TGGGCCACAGAGTGGACAGCTGTCTTGGCGGTTCTCTGTATCATCTCTGGTGCATGSCA 751  
Db 758 CAGGCCACAGAGTGGACAGCTGTCTTGGCGGTTCTCTCTCTATCATCTCTGGTGCATGSCA 817  
Qy 752 CCATGCCCTTTGATGGGCATGACCCATGAATCTCTAGTGAACAGATCAGTAACGGGCTT 811  
Db 818 CCATGCCCTTTGATGGGCATGACCCATGAATCTCTAGTGAACAGATCAGTAACGGGCTT 877  
Qy 812 ACCGTGAGCGCCCAAGCGCTCCGATCGCTGTGGCTGTATCCGCTGGCTGTTAATGGTGA 871  
Db 878 ACCGGAGCGCACCTAAACCTCTGATGCCCTGTGGCTGTATCCGCTGGCTGTGTGATGGTGA 937  
Qy 872 ACCCAACCGCTGGGGCCACACTGGAGGATGTAGCCAGTCAATTTGGTGGGTCAACTGGGGTT 931  
Db 938 ACCCAACCGCGCGGCCACCTCGAGGATGTGGCCAGTCACTGTGGGTCAACTGGGGCT 997  
Qy 932 ACACCACCGAGTCCGGGAACAGGAAGCCCTGCTGAGGCTGGGCAACCTAGTGGTGAAT 991  
Db 998 ACGCCACCGAGTGGGAGAGCAGGAGGCTCCGATGAGGCTGGGCAACCTGGCAGTGAAT 1057  
Qy 992 TTGSCCGGGCTCCATCGCGAGCTGGTTACGTGCTCTCGCGCCCTCTCTGGAGAAATG 1051  
Db 1058 CTGCCCCGGCTCATGCGCTGACTGGCTCGCGCTTCTCCCGCCCTCTCTGGAGAAATG 1117  
Qy 1052 GAGCAAGGTGTGAGCTTCTTCAAGCAGCACTGCGGGAGGTGGGAAGCACTGTACCTG 1111  
Db 1118 GGGCCAAAGTGTGAGCTTCTTCAAGCAGCACTGCGCTGGTGGGGAAGCAACCAACCTG 1177  
Qy 1112 GGCTGGAGCGGCAACATTTCTTAAAGAGTCCCGAAAGGAGAAATGATGCTCAAAATC 1171  
Db 1178 GCCTGGAGCGGCAAGATTCGCTCAAGAAAGTCCCGAAAGGAGAAATGATGCTCAAGTCTC 1237  
Qy 1172 TGCAAGGTGAGCGGCTGAGGATACCTTCTCTCGCCCTGGCAAGAGCAGCTTAAAGCTTC 1231



Db 1238 TCACAGTGACAGCGGTGATGACACTGCCATCGCCCTGGCAAGCAACTCAAGCTGC 1297  
Qy 1232 CGAAGGCATTTCAAGAAAAGTCTCTACCTCGTCAGGGGAGGTACAGAGACCTC 1291  
Db 1298 CAAAGGCGATTTCAAGAAAGGTGTGAGCTCTGCAAGAGGGGTACAGAGACCTC 1357  
Qy 1292 AGGAACCTCAGACCGGTGCTGATCTCAGGAGCGCTGCTGCTGCTGCTGCTGCTGCT 1351  
Db 1358 CGAGCTCAGCCCATCCCTGCGAGCCAGGCGAGGTGCTGCTGCTGCTGCTGCTGCTGCT 1408  
Qy 1352 CAAGAAAGGCATCTTAAGAAAGTCTCGACAGCGTGAATCTGTTACTCTCTCTCCAG 1411  
Db 1409 CCAAGAGGCGATTTCAAGAAAGCCCGACAGCGGAGTCTGGCTACTCTCTCTCCG 1468  
Qy 1412 AGCCAGCGAGTCTGGGGAATCTTAGACGCGAGTGTGTTGTGAGTGGGACCCCG 1471  
Db 1469 AGCCAGTGAATCTGGGAGTCTTGGACGCGAGCGAGTGTGTTGTGAGTGGGATCCCA 1528  
Qy 1472 TGAGACAGAGTCTCCACAGGCTTCAGGCTTCCCTCCACCGCAAGGCAATCTCAAC 1531  
Db 1529 AGGACAGAGCTTCCGAGCTTCAGGCTTCCCTCCATCGCAAGGCAATCTCAAC 1588  
Qy 1532 TCAATGCAAGTCTCCCGCACAGCTTTAGAGGCACTACCCCTAGCACCTTTTGGCTCCC 1591  
Db 1589 TCAATGCAAGTCTCCCGCACAGCTTTAGAGCTTGGAGCTCGGCGCCCCACACCTTCCGCTCCC 1648  
Qy 1592 TGACCAACTGGCTCTCCGATCTCTGAGCGCGCGCCAGCGCCCTCAGGGGCTGTGA 1651  
Db 1649 TGGATGAATCTCGCCACCTCGCCCTCGGCGCGCGCGAGCCCTCAGGGGCTGTGA 1708  
Qy 1652 GTGAGGACAGCTCTGCTCGAGTCTTGAACCAATTGGACTTGCCTGAACTCTTC 1711  
Db 1709 GCGAGGACAGCTCTGCTCGAGTCTTGAACCAATTGGACTTGCCTGAACTCTTC 1768  
Qy 1712 CCGAAACCCCACTGAGGGGCTGTGTCTGTGGAACAACCTGAGGGGCTTGGAGCGCTC 1771  
Db 1769 CAGAGCCCCCACTGCGGGGCTGTGTCTGTGGAACAACCTCAGGGGCTTGGAGCGCTC 1828  
Qy 1772 CTTGAGG-----GTCTGAGGAGTGTGTGAGGAGTGTGTGAGGAGTGTGTGAGGAGTGTGT 1819  
Db 1829 CTTGAGGAGGCTTGGAGGCTGTGTGAGGAGTGTGTGAGGAGTGTGTGAGGAGTGTGT 1888  
Qy 1820 GCTTTTCTGTGACAGCTGCGAAGGCTGTGTGAGGAGTGTGTGAGGAGTGTGTGAGGAGTGTGT 1879  
Db 1889 GCTTTTCTGTGACAGCTGCGAAGGCTGTGTGAGGAGTGTGTGAGGAGTGTGTGAGGAGTGTGT 1948  
Qy 1880 GCTCAAGCTCAGC 1893  
Db 1949 GCTCAAGCTCACC 1962

RESULT 11  
CQ714327  
LOCUS CQ714327 1887 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 261 from Patent WO02068579.  
ACCESSION CQ714327  
VERSION CQ714327.1 GI:42275184  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kits, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 261 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source Location/Qualifiers  
1.1887  
/organism="Homo sapiens"

/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 71.7%; Score 1357; DB 6; Length 1887;  
Best Local Similarity 84.1%; Pred. No. 0;  
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;  
Qy 47 CTCTCGCCCTGGCTCGGAGAGCGCCGCGCTGGCGGAGCGGCTCATCAAGTCGCTTA 106  
Db 35 CCACTCCCTCGGCGCAGAGCTAGCCCGCGCTGGCGGAGGCTCATCAAGTCGCTTA 94  
Qy 107 AACCTCTGATGAAGAACGAGCGGTGAAGCGGACCATCAAAACACAACTCTGGCGACC 166  
Db 95 AGCCCTTAATGAGAGCAGCGGTGAAGCGGACCAACACAGCAGCAAACTCTGGCGACC 154  
Qy 167 GCTACGAGTTCTTGGAGA CCGTGGGCAAGGCGCCTACGGAAGGTGAAGAAGCAGCAG 226  
Db 155 GCTACGAGTTCTTGGAGACCTTGGGCAAGGCACTACGGAAGGTGAAGAAGCGCGGG 214  
Qy 227 AGAGCTCGGCGCTCTGGTGGCCATCAAGTCCATCAGGAAGAGCAAAATCAAGATGAGC 286  
Db 215 AGAGCTCGGCGCGCTCTGGTGGCCATCAAGTCAATCCGGAAGGACAAAATCAAGATGAGC 274  
Qy 287 AGGATCTGTGCACATACGAGGAGGAGATTGAGATCATGTCTCACTCAACCCCTCACA 346  
Db 275 AGGATCTGTGCACATACGAGGAGGAGATTGAGATCATGTCTCACTCAACCCCTCACA 334  
Qy 347 TCATTGCTCATCAAGTGTGAGAAATGAGCAAGATGTGATGTGATGATGATGATGATG 406  
Db 335 TCATTGCTCATCAAGTGTGAGAAATGAGCAAGATGTGATGTGATGATGATGATGATG 394  
Qy 407 CAGCCGAGGAGTCTGTATGATTAATCATGAGTGGCGCCACCGCTGAGTGGCGGAGC 466  
Db 395 CAGCCGAGGAGTCTGTATGATTAATCATGAGTGGCGCCACCGCTGAGTGGCGGAGC 454  
Qy 467 CAGGCAATTTCTCCGACAGATCGTGTCTCGCTGCACTACTGCGACAGAACCGGATCG 526  
Db 455 CTAGGCATTTCTCCGCGAGTGTCTCTCGCGTGCATATTGCCATCAGACAGAGTTG 514  
Qy 527 TTCAACGAGATCTCAAGCTGGAAAAACATCTCTTAGATGCCAATGGAAAAATCAAGATTG 586  
Db 515 TCACACGAGATCTCAAGCTGGAGAACATCTCTTTGGATGCCAATGGGAATATCAAGATTG 574  
Qy 587 CTGACTTTGGCTCTCCAACTGTACCAAAAGCAGTTCTTCCACAGCTTCTGTGGGA 646  
Db 575 CTGACTTTGGCTCTCCAACTGTACCAATCAAGCAAGTTCTTCCAGACATTTCTGTGGGA 634  
Qy 647 GCGCTCTCTACGCTCGCCTGAGATAGTCAACGGAAGCCCTATGTGGGCGCAGAGTGG 706  
Db 635 GCGCTCTCTATGCTCGCCAGAGATTGTCAATGGGAAGCCCTACACAGGCCAGAGTGG 694  
Qy 707 ACAGCTGCTCTGTGGCGTTCTCTGTACATCTCTGTGATGAGCAATGCAATGCTTTGACG 766  
Db 695 ACAGCTGCTCTGTGGCGTTCTCTGTACATCTCTGTGATGAGCAATGCAATGCTTTGATG 754  
Qy 767 GGCAGGATCAATAAACAACCTGTGAGCAAAATCAGTACGGGGCTTACCGTGGCGGCCA 826  
Db 755 GGCATGACCAATAAGATCTTAGTGAACACAGATCAGCAACGGGGCTTACCGGAGGCCACTTA 814  
Qy 827 AGCGCTCCGATGCTGTGGCTGTATCGGTGGCTGTTAATGGTGAACCCACCGCTGGGG 886  
Db 815 AACCTCTGATGCTGTGGCTGTATCGGTGGCTGTTGATGGTGAACCCACCGCTGGGG 874  
Qy 887 CCACACTGAGAGGATGATAGCAGTCAATGTTGGTGGTCAACTGGGGTTACACACCGGAGTGC 946  
Db 875 CCACCTTGGAGGATGTGGCGAGTCACTGTGTGGTCAACTGGGGTGTACGCCACCGGAGTGG 934  
Qy 947 GGGAAACAGAGGAGCCCTGCGTGGGTTGGGCACTCTAGTGTGATCTTTGGCGGCGCTCCA 1006  
Db 935 GAGAGCAGGAGGCTCCCGCATGAGGGTGGGCACTCTGCGAGTGTGCTCTGCGCGCTTCCA 994  
Qy 1007 TGGCGGACTGTTACGTGCTCTCTCGCGCCCTCTCTGAGAAATGGAGCAAGGTGTGCA 1066



Db 995 TGGCTGACTGCTCCGGCGTTCTCCCGCCCTCTGGAGAAATGGGCCCAAGGTGTGCA 1054  
Qy 1067 GCTTCTTCAAGCAGCAGCTGCGGAGGTGAAGCAGTGTACTGCTGGGCTGGAGCGGCAAC 1126  
Db 1055 GCTTCTTCAAGCAGCAGTCACTGCTGGGGAAGCACCACCCCTGGCTGGAGCGCCAGC 1114  
Qy 1127 ATTCTTCTAAGAGTCCGAAAGGAGATGATGGCTCAAAATCTGCAAGGTGACCCGG 1186  
Db 1115 ATTCTTCAAGAGTCCGCAAGGAGATGATGGCTCAAAATCTGCAAGGTGACCCGG 1174  
Qy 1187 CTGAGGATACCTTCTGCTCCCTGCGAGAGCAGCTTAACTTCTCGAAGGCAATTTCTCA 1246  
Db 1175 CTGATGACACTGCTCCCTGCGAGAGCAACCTCAAGCTGCCAAGGCAATTTCTCA 1234  
Qy 1247 AGAAAGAGTCTCTACCTCTGTCAGGGAGGTACAGGAGGACCTCAGGAACTCAGACCGG 1306  
Db 1235 AGAAGAGGTGTAGCTCTCAGAGGGGTACAGGAGGACCTCAGGAGCTCAGCCCAA 1294  
Qy 1307 TGCCTGATATCTCAGGGCAGCTGTCTGCTGTATCCCTGCTCCCAAGGAAAGGCAATCC 1366  
Db 1295 TCCCTGCGAGCCAGGCGAGCTGCC-----CCCTGCTCCCAAGGAGGCAATTC 1345  
Qy 1367 TTAAGAGTCTGACAGGCTGAATCTGGTTACTACTCTCTCTCAGAGCCAGCAGTCTG 1426  
Db 1346 TCAAGAGCCCGCAGCAGCGAGTCTGGCTACTACTCTCTCTCCGAGCCAGTGAATCTG 1405  
Qy 1427 GGGAACTCTTAGAGCCAGTGTCTGTTGAGTGGGACCCCTGCGAGCAGAGTCTC 1486  
Db 1406 GGGAGCTCTTGAGCGAGCGAGCTGTTGAGTGGGATCCCAAGAGCAGAGGCTC 1465  
Qy 1487 CACAGGCTTCAGGCTCTCTCTCCACCGCAGGCAATCTCAAACTCAATGGCAAGTTCT 1546  
Db 1466 CGCAAGCTTCAGGCTCTCTCTCCATCGCAAGGCAATCTCAAACTCAATGGCAAGTTCT 1525  
Qy 1547 CCAGCAGCTTGAAGGCACTACCTTACCTTGGCTTCTCTGAGCCAGCAGTCTG 1606  
Db 1526 CCGAGCAGCTTGGAGCTCGCGCCCGCCACCACTCTCGGCTCGATGAACTCGCC 1585  
Qy 1607 CTTCCACTCTGAGCGCGCCAGCCCGCTCAGGGGCTGTGAGGAGCAGCAGTCC 1666  
Db 1586 CACTCGCCCTGCGCGCGCCAGCCAGCCCTCAGGGGCTGTGAGGAGCAGCAGTCC 1645  
Qy 1667 TGTCTCGAGTCTCTTTGACCAATTTGCACTTGGCTTGGCTTCTCCGAAACCCCACTGA 1726  
Db 1646 TGTCTCTGAGTCTTTGACCACTTGGCTTGGCTTGGCTTCTCCGAGCCCGCCACTGC 1705  
Qy 1727 GGGCTGTGTCTGTGGAACCTGAGGGGCTTGGAGGCTCTCTCTGAAAG----- 1780  
Db 1706 GGGCTGTGTCTGTGGAACCTGAGGGGCTTGGAGGCTCTCTCTGAAAG----- 1765  
Qy 1781 -----GTCTGAGCGATGGGCGAGGAACTCTTGGGGGATAGTCTTTCTCTGACAG 1834  
Db 1766 GAGCTGCTGAGGCTGTGGCGGAGATCTTTGGGGGAGAGTCTTTTCTCTGACAG 1825  
Qy 1835 ACTGCCAAGAGGTGACTGCAAGCTTACAGCAAGCCCTAGGAATCTGCTCAAGCTCAGC 1893  
Db 1826 ACTGCCAAGAGGTGACAGGAGCTTACCGAGGCACTGAGGGTCTGCTCAAGGCTCACC 1884

RESULT 12

Q0782778 3395 bp DNA linear PAT 17-MAR-2004  
LOCUS Sequence 2918 from Patent EP1396543.  
DEFINITION Q0782778  
ACCESSION Q0782778  
VERSION Q0782778.1 GI:45502721  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,

Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.  
Primers for synthesizing full length cDNA clones and their use  
Patent: EP 1396543-A 2918 10-MAR-2004;  
Research Association for Biotechnology (JP)  
Location/Qualifiers  
source 1. 3395  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
118. 2004  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAF86008.1"  
/db\_xref="GI:45502722"

CDS

/translation="MESLVEARRSGPTPSAAELARPLAELGLIKSPKPLMKQAVKRHH  
KHNLRHRYFLETGRTGSKVKKARESSRLVAIKSIRDKIKDSRDLHRRLE  
IMSLNHPHIIAIEHPENSISKVIYNEIASRGDLYDIISERQQLSQQDLHRRLE  
SAVYQHNRVVRHDKLENILLDANGINKIADFGLSNLYHQKFLQTCGSPLYASP  
EIVNGKPYTPGVDSWISGLVLLVILVHTMPFDGHDHKLIVKQISNAYRPPKPSAD  
CGLIRMLLVNPTTRATLEDAVSHVMWCVATRVGEQAEHPHGGHPSDSARASMD  
WLRSSRPILLEAKVCFFQHPAPGGSTTPGLERHSLKSKRKNDAQSLHSDTA  
DDTAHRPKNLKPLKGLIKKVASAAGVQEDPPPELSPASPQAPKLPKGLIK  
KPRQESGYTSPSPESBELLDAGDVFSGDPKQPKQASGLLHRKILKILKNGLP  
SOTALELAAPTTFGSLDELAPRPLARSPGSAVSDSILSSFPQLDLPRLPEP  
PLRGCVSDNLTLGLEPPSPGSGCLRRWRQDPLGDSFCSLTDQCEVTATVQALRVC  
SKLT"

ORIGIN

Query Match 71.7%; Score 1357; DB 6; Length 3395;  
Best Local Similarity 84.1%; Pred. No. 0;  
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;  
Qy 47 CCTCGCGCTGGCTCGAGAGCGCCGCGCGCTGGCGGAGCGGCTCATCAAGTCGCCTA 106  
Db 152 CCATCTCCTCGCGCGAGAGCTAGCCGCGCGCTGGCGGAGGCGCTCATCAAGTCGCCA 211  
Qy 107 AACCTCTGATGAAGAAGCAGCGGTGAAGCGGACCATCAAAACACAACTCTGGCGACC 166  
Db 212 AGCCCTAATGAAGAAGCAGCGGTGAAGCGGACCATCAAAACACAACTCTGGCGACC 271  
Qy 167 GTACGAGTTCCTGGAGACCTGGGCAAGGCGACCTACGGGAAGGTGAAGAAGCAGCAG 226  
Db 272 GCTACGAGTTCCTGGAGACCTGGGCAAGGCGACCTACGGGAAGGTGAAGAAGCAGCAG 331  
Qy 227 AGAGCTCGGCGCTCTGGTGGCCATCAAGTCCATCAGGAAGAAGCAAAATCAAGATGAGC 286  
Db 332 AGAGCTCGGCGCGCTGGTGGCCATCAAGTCCATCAGGAAGAAGCAAAATCAAGATGAGC 391  
Qy 287 AGGATCTGCTGCATACAGGAGGAGATTGAGATCATGTCTTCACTCAACCCACCCACA 346  
Db 392 AGATCTGATGCATACAGGAGGAGATTGAGATCATGTCTCATCTCAACCCCTCACA 451  
Qy 347 TCATTGCTATCCATGAAGTGTGTTGAGAAATAGCAGCAAGATTGTGATGTCTCATGAGATG 406  
Db 452 TCATTGCTATCCATGAAGTGTGTTGAGAAATAGCAGCAAGATTGTGATGTCTCATGAGATG 511  
Qy 407 CGAGCGAGGCGATCTGTATGATTACATCAGTACGCGGCCACCGCTGAGTACGCGGAGC 466  
Db 512 CCAGCGGCGCGACCTTTTATGACTACATCAGCAGCGGCGAGCAGCTCAGTACGCGGAG 571  
Qy 467 CCAGGCAATTTCTCCGACAGATCGTGTCTCCCTGCACTACTGCGCACAGAACCGGATCG 526  
Db 572 CTAGGCATTTCTCCGCGAGATCGTCTCTCGCGTGCATATTGCGCATCAGACAGAGTTG 631  
Qy 527 TTCACCGAGATCTCAAGCTGGAAAAACATCTCTTAGATGCCAATGGAAACATCAAGATTG 586  
Db 632 TCCACCGAGATCTCAAGCTGGAGAAATCTCTTTGGATGCCAATGGAAATATCAAGATTG 691  
Qy 587 CTGACTTTGGCCTCTCCAACTGTACCAAAAGCAAGTTCTCTCCAGACGTTCTGTGGGA 646  
Db 692 CTGACTTTGGCCTCTCCAACTGTACCAATCAAGCAAGTTCTCTCCAGACATTTCTGTGGGA 751



```
QY 287 AGGATCTGCTGCACATACGGAGGAGATTGAGATCATGTCTTCTCACTCAACACCCCCACA 346
Db 392 AAGATCTGATGACATACGGAGGAGATTGAGATCATGTCTCACTCAACACCCCTCACA 451
QY 347 TCATTGCCATCCATGAAGTGTTCAGAAATAGCACCAAGATTGTGATTTGTCATGAGTATG 406
Db 452 TCATTGCCATCCATGAAGTGTTCAGAAACAGCAAGATCGTGTATGTCATGAGTATG 511
QY 407 CCAGCGAGGGGATCTGTATGATTACATCAGTGAAGCGCCACCGCTGAGTGAGCGGACG 466
Db 512 CCAGCGGGGCGACCTTTATGACTATACATCAGCAGCGGACGACGCTCAGTGAGCGCGAAG 571
QY 467 CCAGGCAATTTCTCCGACAGATCGTGTCTGCCCTGCACCTACTGCGCACAGAACGGGATCG 526
Db 572 CTAGGCAATTTCTCCGGCAGATCGTCTCTCGCGTGCACTATTGCCATCAGAAACAGAGTTG 631
QY 527 TTACCGGAGATCTCAAGCTGGAACATCCTTTCTAGATGCCAATGGAAACATCAAGATTG 586
Db 632 TCCACCGGAGATCTCAAGCTGGAGAACATCCTCTTTGGATGCCAATGGGAATATCAAGATTG 691
QY 587 CTGACTTTGGGCTCTCCAACTGTACCAAGCAAGTTCTCTCCAGACGTTCTGTGGGA 646
Db 692 CTGACTTGGGCTCTCCAACTGTACCAATGAAGCAAGTTCTCTCCAGACATTTCTGTGGGA 751
QY 647 GCCCTCTCTACGCTCGCCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCGACAGGTGG 706
Db 752 GCCCTCTCTATGCTCGCCAGAGATTGTCAATGGGAAGCCCTACACAGGCCCGACAGGTGG 811
QY 707 ACAGTGTCTCTGGGGCTTCTCTGTATCATCTCTGTGTGATGGCAATGCGCATGCGCTTTGACG 766
Db 812 ACAGTGTCTCTGGGGCTTCTCTGTATCATCTCTGTGTGATGGCAATGCGCATGCGCTTTGATG 871
QY 767 GGCAGGATCAAAACACTGCTGAAGCAAACTACAGGGGCTTACCGTGAAGCGGCCA 826
Db 872 GGCATGACCAATAAGATCTCTAGTGAACAGATCAGAAACGGGGCTTACCGGAGGCACTTA 931
QY 827 AGCCCTCGATGCTGCGCTGTATCGGCTGGCTGTAAATGTTGAACCCACCCCGTGGG 886
Db 932 AACCTCTGTATGCTGTGGCTGTATCGGCTGGCTGTGATGTTGAACCCACCCCGCGGG 991
QY 887 CCACACTGGAGGATGTAGCCAGTCAATTTGGTGGGTCAACTGGGGTTACACACCGGAGTGG 946
Db 992 CCACCTCTGGAGGATGTGGCCAGTCACTGGTGGGTCAACTGGGGCTACGCCACCCGAGTGG 1051
QY 947 GGGAAACAGAGCCCTGCTGAGGCTGGCACCTAGTGTGATCTTTGGCGGGCCCTCCA 1006
Db 1052 GAGAGCAGAGGCTCCGATGAGGCTGGCACCTGCGGAGTGTCTGCGCGCGCCCTCCA 1111
QY 1007 TGGCGGACTGTTTACGTCTGCTCGCGCCCTCTCGGAGAAATGGAGCCAAAGTGTGCA 1066
Db 1112 TGGCTGACTGCTCGGGGTTCTCCCGCCCTCTCTGGAGAAATGGGGCCAAAGTGTGCA 1171
QY 1067 GCTTCTTCAAGCAGACGTGCGGAGAGTGAAGCACTGTACTCTGGGTGGAGCGGCAAC 1126
Db 1172 GCTTCTTCAAGCAGCATGCTGCTGGTGGGGAAGCACCACCCCTGGCTGGAGCGCCAGC 1231
QY 1127 ATTCTCTTAAGAGTCCGAAGGAGATGATGCTCAAAATCTGCAAGGTGACCCGG 1186
Db 1232 ATTGCTCAAGAGTCCCGAAGGAGATGATGATGAGCCAGTCTCTCCACAGTACACGG 1291
QY 1187 CTGAGGATACCTCTTCTCGCCCTGGCAAGACAGCGCTTAAAGCTTCCGAAAGGCAATTTCA 1246
Db 1292 CTGATGACATGCCCCATGCCCCCTGGCAGAGCAACCTCAAGTGGCAAGGGCAATTTCA 1351
QY 1247 AGAAAAAGTCTCTTACTCTGTCAGGGAGGTACAGGAGGACCCCTCAGGAACTCAGACCGG 1306
Db 1352 AGAAGAAGGTGTACGCTCTGAGAAAGGGTACAGGAGGACCCCTCCGAGCTCAGCCCAA 1411
QY 1307 TGCCTGATCTCCAGGGAGGCTGCTGCTGTGTATCCCTGCTCTCCAGGAAAGGCATCC 1366
Db 1412 TCCCTGCGAGCCAGGGAGGCTGCC-----CCCTGCTCTCCCAAGAGGGCATTC 1462
QY 1367 TTAAGAGTCTCGACAGCGTGAAATCTGTGTTACTTCTCTCCAGCGCCCGAGGCTGTG 1426
```

```
Db 1463 TCAGAAGACCCCGACAGCGGAGTCTGGCTACTACTCTCTCCGAGCCCGAGTGAATCTG 1522
QY 1427 GGGAATCTTTAGAGCGCAGTGATGTTTGTGTAGTGGGACCCCGTGGAGCAGAAGTCTC 1486
Db 1523 GGAGAGCTTTTGGAGCGCAGGCGAGCTGTTTGTGTAGTGGGATGCCAAGGAGCAGAAGCTC 1582
QY 1487 CACAGGCTTCAGGGCTCTCTCTCCACCGCAAGGCAATTTCTCAAACTCAATGGCAAGTTCT 1546
Db 1583 CGCAAGCTTCAGGGCTCTCTCTCCATCGCAAGGCAATTTCTCAAACTCAATGGCAAGTTCT 1642
QY 1547 CCAGCAGAGCTTTAGAAGGCACTACCCCTTAGCCTTTTGGCTCTCTGGGCAAACTGGCCT 1606
Db 1643 CCAGACAGAGCTTTGGAGCTCGCGGCCCCACCACTTGGCTCTCTGGATGAATCTCGCCC 1702
QY 1607 CTTCCCATCTCTGACGCGCGCCCGCCAGCGCCCTCAGGGGCTGTGAGTGAGCAGCAGCATCC 1666
Db 1703 CACTCGCCCCCTTGGCCCGGGCCAGCGGACCTCAGGGGCTGTGAGGAGGAGCAGCATCC 1762
QY 1667 TGTCTCTCCGAGTCTCTTTGACCAATTTGGACTTGGCTTGAACGCTTCCCGAAACCCCACTGA 1726
Db 1763 TGTCTCTGAGTCTCTTTGACCACTGGACTTGGCTTGAACGCTTCCCGAGAGCCCACTGC 1822
QY 1727 GGGGCTGTGTCTGTGGCAAACTCAGGGGGCTTGAAGAGCCTCCCTCAGAAAG----- 1780
Db 1823 GGGGCTGTGTCTGTGGCAAACTCAGGGGGCTTGAAGAGCCTCCAGAGGCGCCCTCAGAGGGCCCTG 1882
QY 1781 -----GTCTGAAGCAGTGTGTGGAGGAATCTTGGGGATAGCTGCTTTCTCTGACAG 1834
Db 1883 GAAGCTGCTTGGAGCGCTGGCGCAGGATCTCTTTGGGGGACAGCTGCTTTTCCCTGACAG 1942
QY 1835 ACTGCCAAGAGGTGACTGTCAGCCTTACAGACAAGCCCTAGGAAATCTGTCTCAAAGCTCAGC 1893
Db 1943 ACTGCCAGGAGGTGACAGCGACCTTACCGACAGCAGCTGAGGGTCTGTCTCAAAGCTCACC 2001

RESULT 14
AK074830
LOCUS
DEFINITION
Homo sapiens cDNA FLJ90349 fis, clone NT2RP2003179, weakly similar
to PUTATIVE SERINE/THREONINE-PROTEIN KINASE EMK (EC 2.7.1.1.-).
ACCESSION
AK074830
VERSION
AK074830.1 GI:22760531
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
AUTHORS
Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,
Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,
Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
Kojima, S., Nagahara, K., Masuho, Y., Ono, T., Okano, K., Yoshikawa, Y.,
Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
Ninomiya, K.
TITLE
NEDO human cDNA sequencing project
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 3395)
AUTHORS
Isogai, T. and Otsuki, T.
TITLE
Direct Submission
JOURNAL
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel.81-438-52-3975, Fax:81-438-52-3986)
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
FEATURES
source
1. .3395
/organism="Homo sapiens"
```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2003179"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/notes="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction"
118..2004
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAC11234.1"
/db_xref="GI:22760532"
/translation="MESLVFARRSGPTPSAAELARPLAELIKSPFLMKKQAVKRHH
HKNNRHYRFLGTGTYGKVKKARESSRLVAIKSRDKIKDQDLMIHRREIE
IMSLNHPIHIAHEVPENSKIIVNMEYASRGDLDYISERQOLSEARHPFRQIV
SAVHYCHONVVRHDLKLENIKIADFNGLNLYHQKFLQTFCGSPLYASP
EIVNGKPYTGTPGWSGLGVLVYLVHGTWPFDFGHKLVKIOISNGAYRPPKPSDA
CGLIRWLLVNPTRRATLEVDVASHWVWGYATFVGEQAPHEGSHGSPSARASND
WLRSSRPILLENGKVCSPFKQHPAGGSGTTPGLERHSLKSRKENDMAQSLHSDTA
DTRAHREKNNLKPGLIKLKKVSAASAEVQEDPPELSPIPASFGQAPLPPKKGLIK
KPRQREGSYSSPEPSGELLDAGDFVSGDPKQKPPQASGLLHRKILKLNKGF
SOTALELAAPTTFGSDLELAPPPLARASRPSGAVSDESLISSEFQDLDPRLPEP
PLRGCVSDNLTLGLEBPSPGSGSLRRWRQDPLGDSFSLTDQEVTVATYRQALRVC
SKLT"
Query Match 71.7%; Score 1357; DB 9; Length 3395;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;
QY 47 CCTCCGCTGGCTCGAGAGCCCGCGCGTGGCGGAGCGGCTCATCAAGTCGCTA 106
DB 152 CCATCTCCCGCGCAGAGCTAGCCCGCGCGTGGCGGAGGCGTCAAGTCGCTA 211
QY 107 AACCTCTGATGAAGAAGCAGCGGTGAAGCGGACCATCAAAACAACTCGCGGACC 166
DB 212 AGCCCTAATGAGACGAGCGGTGAAGCGGACCATCAACAGCACAACTCGCGGACC 271
QY 167 GTACAGATTCCTGGAGACGCTGGGCAAGGACCTACGGAAGGTGAAGAAGCACGAG 226
DB 272 GTACAGATTCCTGGAGACGCTGGGCAAGGACCTACGGAAGGTGAAGAAGCGCGG 331
QY 227 AGAGCTCGGCGGCTCGTGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGC 286
DB 332 AGAGCTCGGCGGCTCGTGCCATCAAGTCCATCAGGAAGGACAAATCAAGATGAGC 391
QY 287 AGGATCTGCTGCATACGAGGAGGATGAGATCATGTCTTCACTCAACACCCCCACA 346
DB 392 AAGATCTGATGCATACGAGGAGGATGAGATCATGTCTCACTCAACACCCCTACA 451
QY 347 TCATTGCCATCCATGAAGTGTGAGAAATAGCAGCAAGATTGATGATGATGATGATG 406
DB 452 TCATTGCCATCCATGAAGTGTGAGAAATAGCAGCAAGATTGATGATGATGATGATG 511
QY 407 CCAGCCGAGCGATCTCTATGATTACATCAGTAGCGGCGCAAGTGTGAGTGAGCGG 466
DB 512 CCAGCCGAGCGATCTCTATGATTACATCAGTAGCGGCGCAAGTGTGAGTGAGCGG 571
QY 467 CCAGGCAATTTCTCCGACAGATCGTGTCTGCGCTGCACTACTGCGACAGAAAGGATCG 526
DB 572 CTAGGCATTTCTCCGCGCAGATCGTCTCTGCGGTGCACTATGCGCATCAGAACAGAGTTG 631
QY 527 TTACCGAGATCTCAAGCTGGAACATCTCTTCTAGATGCAATGGAAACATCAAGATTG 586
DB 632 TCACCGAGATCTCAAGCTGGAACATCTCTTCTAGATGCAATGGAAACATCAAGATTG 691
QY 587 CTGACTTTGGCTCTCCAACTGTACCAAGCAAGCAAGTTCTTCCAGACGCTCTGTGGGA 646
DB 692 CTGACTTTGGCTCTCCAACTGTACCACTGTACCACTAGGCAAGTTCTTCCAGACATCTGTGGGA 751
QY 647 GCCTCTCTACGCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCAGAGGTGG 706

```

ORIGIN

```

DB 752 GCCCCTCTATGCTGCCAGAGATTGTCAATGGGAAGCCCTACACAGGCCCCAGAGTGG 811
QY 707 ACAGCTGGTCTCTGGCGGTTCCTCTGTACATCTCTGTGTCATGCCACCATGCTTTGACG 766
DB 812 ACAGCTGGTCCCTGGGTGTTCTCTCTATACATCTCTGTGTCATGCCACCATGCTTTGATG 871
QY 767 GGCAGGATCATAAACACTGGTGAAGCAATCATGTAACGGGGCTTACCTGAGCGGCCCA 826
DB 872 GGCATGACCATTAAGATCTTAGTGAACACGGGGCTTACCGGGAGCCACCTA 931
QY 827 AGCCGTCGATGCTCTGGCTCATCCGGTGGCTGTAAATGTTGAACCCACCCCTCGGG 886
DB 932 AACCTCTGATGCTCTGGCTCATCCGGTGGCTGTAAATGTTGAACCCACCCCGGG 991
QY 887 CCACATCGAGGATGTAGCCAGTCAATTTGGTGGTCACTCTGGGTTTACACACCGGAGTGG 946
DB 992 CCACCTCTGGAGGATGTGGCCAGTCACTGGTGGTCACTCTGGGCTACGCCACCCGAGTGG 1051
QY 947 GGGAA CAGGAAGCCCTCGCTGAGGCTGGGACCCCTAGTGGTGAATTTGGCGGGCTTCCA 1006
DB 1052 GAGAGCAGGAGGCTCCGCATGAGGGTGGGCAACCCTGGCAGTGAATCTGCCCCGGCTTCCA 1111
QY 1007 TGGCGGACTGGTTACGCTCGCTCTCGCGGCCCTCTCTGGAGAAATGGAGCCAAAGTGTGCA 1066
DB 1112 TGGCTGACTGGCTCGCGGCTTCTCTCGGCCCTCTCTGGAGAAATGGGCGCAAGGTGTGCA 1171
QY 1067 GCTTCTTCAAGCAGCAGTGGCGGAGGTGGAAAGCACTGTATCTGGGCTGGAGCGGCAAC 1126
DB 1172 GCTTCTTCAAGCAGCAGTGCACCTGGTGGGGAAGCACCAACCCCTGGGCTGGAGCGGCGAG 1231
QY 1127 ATTCTCTTAAGAGTCCCGAAGGAGAAATGACATGCTCAAAATCTGCAAGGTGACCCGG 1186
DB 1232 ATTCTCTTAAGAGTCCCGAAGGAGAAATGACATGCTCAAAATCTGCAAGGTGACCCGG 1291
QY 1187 CTGAGGATACCTTCTCTCGCCCTGGCAAGCAGCAGCTTAAGCTTCCGAAAGGATCTCTCA 1246
DB 1292 CTGATGACACTGCCCCATCGCCCTGGCAAGCAGCACTCAAGCTGCCAAGGGGATCTCTCA 1351
QY 1247 AGAAAGTCTCTTACTCTGTCTGAGGGAGGTACAGAGGACCTCTAGGAATCTCAGACCGG 1306
DB 1352 AGAAGAGGTGTGAGCTCTGCAAGAAAGGGGTACAGAGGACCTCTCGGAGCTCAGGCCAA 1411
QY 1307 TGCTCTATCTCCAGGCGAGCTGCTCCCTGTGTATCTCTGCTCCCAAGAAAGGATCTC 1366
DB 1412 TCCTCGGAGCCCGAGGCGAGCTGCCCC-----CCCTGCTCCCAAGAGGAGGATTC 1462
QY 1367 TTAAGAAGTCTCGACAGCGTGAATCTGTGTTACTTCTCTCCAGAGCCCGAGGCTGTG 1426
DB 1463 TCAAGAAGCCCGAGCAGCGAGTCTGGCTACTTCTCTCCGAGCCCGAGTGAATCTG 1522
QY 1427 GGGAACTCTTAGCGCCAGTGTGTTTGTGAGTGGGACCCCGTGGAGCAGAGTCTC 1486
DB 1523 GGGAGCTCTTGGACGCGAGCGAGCTGTTTGTGAGTGGGATCCCAAGGAGCAGAGCCCTC 1582
QY 1487 CACAGCTTTCAGGGCTCTCTCCACCGCAAGGAGGATCTCTCAAACTCAATGGCAAGTCT 1546
DB 1583 CGCAAGCTTTCAGGGCTCTCTCCATCGCAAGGAGGATCTCTCAAACTCAATGGCAAGTCT 1642
QY 1547 CCGCA CAGCTTTCAGAGGACCTTACCCCTTAGCACTTTGGCTCTCTTGGACCACTGGCCT 1606
DB 1643 CCAGACAGCTTGGAGCTCGCGGCCCCCAACCTTTCGGCTCTCTCTGGATGAACTCGCCC 1702
QY 1607 CCTCCCATCTCGAGCCCGGCCCGCCCGAGCGGCTCTCAGGGCTGTGATGAGGACAGATCC 1666
DB 1703 CACCTCGCCCTCTGGCGGCCCGGCGAGCCCTCTCAGGGCTGTGAGGAGGAGCAGATCC 1762
QY 1667 TGTCTCTCGAGTCTCTTGTGACCAATTGGACTTGTGCTGAACCTTCCCGAAACCCCACTGA 1726
DB 1763 TGTCTCTCGAGTCTCTTGTGACAGCTGGAATTTGCTGAAAGGCTTCCAGAGAGCCCACTGC 1822
QY 1727 GGGGCTGTGTGTGTGTGGAACAACCTGAGGGGCTTGAAGAGGCTCTCCTCAGAAAG----- 1780

```

CDS

Db 1823 GGGGCTGTGTCTGTGTGACAACTCTCAGCGGGCTTGTAGGAGCCCCCTCAGAGGGCCCTG 1882  
Qy 1781 -----GTCTGAAGCATGTGGCAGGAATCTTTGGGGGATAGCTGCTTTTCTCTGACAG 1834  
Db 1883 GAAGCTGCTGAGCGCTGGCGGACAGATCTTTGGGGGACAGCTGCTTTTCTCTGACAG 1942  
Qy 1835 ACTGCCAAGAGTGTGACGCTTACAGACAGCCCTAGGAATGTCTCAAGCTCAGC 1893  
Db 1943 ACTGCCAAGAGTGTGACGCGACCTACCGAGGACCTGAGGGTGTCTCAAGAGCTCACC 2001

## RESULT 15

BC017306 3402 bp mRNA linear PRI 29-JUN-2004  
LOCUS Homo sapiens likely ortholog of rat SNF1/AMP-activated protein  
DEFINITION kinase, mRNA (cDNA clone MGC:29812 IMAGE:5088037), complete cds.

ACCESSION BC017306.2 GI:33878200

VERSION MGC.

KEYWORDS Homo sapiens (human)

SOURCE MGC.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3402)

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahney, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalios, D.E.,

Schmerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 3402)

Strausberg, R.

Direct Submission

Submitted (05-NOV-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Aug 19, 2003 this sequence version replaced gi:16878205.

Contact: MGC help desk

Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)

Tissue Procurement: Lou Straudt

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: [nisc\\_mgc@nri.nih.gov](mailto:nisc_mgc@nri.nih.gov)

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Teague, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 39 Row: 0 Column: 12  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 13569921.

## FEATURES

## source

1. .3402  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:29812 IMAGE:5088037"  
/tissue\_type="Lymph, lymphoma"  
/clone\_lib="NIH MGC\_99"  
/lab\_host="DH10B"  
/note="Vector: pOTB7"  
1. .3402  
/gene="SNARK"  
/note="synonym: DKFZP434J037"  
/db\_xref="LocustID:81788"  
100. 1986  
/db\_xref="MIM:608131"  
/gene="SNARK"  
/codon\_start=1  
/product="hypothetical protein DKFZp434J037"  
/protein\_id="AAH17306.1"  
/db\_xref="GI:16878206"  
/db\_xref="LocustID:81788"  
/db\_xref="MIM:608131"  
/translation="MESLFAARRSGTPSAELARPLAELGIKSPKPLMKKQAVKRHH  
KNSLHRRHYFLETGLGTYGVKVAKSSGRVLAISKIKOKIDQDMHRRREIE  
IMSLNHPHIIAIEHFVENSSTKIVMEYASRGDLYISERQQLSERAEHFFRQIV  
SAHYCHQNRVHRDLKLENLDANGNKIADFGSLNLYHQGKFLQTCGSPLYASP  
EIVNGKPYTGPEVDSLSGLVLLVILVHGTMPDGHDKILVKOISNGAYREPCKPSDA  
CGILRWLLMWNPRRATLEDVASHVWVWNGYATRVGEQEPHGGHGPSASASMD  
WLKRSRPLLENGAKVCSFFKQHPAGGGSTTPGLERQHSLSKSRKENDMAQSLSHSDTA  
DDTAHPRGKSNLKPGLIKKVSASAEQVDEPPELSPISPASQAPALPKKGLIK  
KPRQRESGYTSPSESGELDAGVDFVSGDPEKQPKQASGLLHRRKILKNGKF  
SOTALBLAAYPTFGSLDELAPPRPARSRPSGAVESDLSSESFDQLDLPLRLPBP  
PLRGCVSDNLTGLEPPEPSEPGSGCLRRWRQDFLGDSCFSLTDCQEVATYRQALRVC  
SKLT"

## CDS

1. .3402  
/gene="SNARK"  
/note="synonym: DKFZP434J037"  
/db\_xref="LocustID:81788"  
100. 1986  
/db\_xref="MIM:608131"  
/gene="SNARK"  
/codon\_start=1  
/product="hypothetical protein DKFZp434J037"  
/protein\_id="AAH17306.1"  
/db\_xref="GI:16878206"  
/db\_xref="LocustID:81788"  
/db\_xref="MIM:608131"  
/translation="MESLFAARRSGTPSAELARPLAELGIKSPKPLMKKQAVKRHH  
KNSLHRRHYFLETGLGTYGVKVAKSSGRVLAISKIKOKIDQDMHRRREIE  
IMSLNHPHIIAIEHFVENSSTKIVMEYASRGDLYISERQQLSERAEHFFRQIV  
SAHYCHQNRVHRDLKLENLDANGNKIADFGSLNLYHQGKFLQTCGSPLYASP  
EIVNGKPYTGPEVDSLSGLVLLVILVHGTMPDGHDKILVKOISNGAYREPCKPSDA  
CGILRWLLMWNPRRATLEDVASHVWVWNGYATRVGEQEPHGGHGPSASASMD  
WLKRSRPLLENGAKVCSFFKQHPAGGGSTTPGLERQHSLSKSRKENDMAQSLSHSDTA  
DDTAHPRGKSNLKPGLIKKVSASAEQVDEPPELSPISPASQAPALPKKGLIK  
KPRQRESGYTSPSESGELDAGVDFVSGDPEKQPKQASGLLHRRKILKNGKF  
SOTALBLAAYPTFGSLDELAPPRPARSRPSGAVESDLSSESFDQLDLPLRLPBP  
PLRGCVSDNLTGLEPPEPSEPGSGCLRRWRQDFLGDSCFSLTDCQEVATYRQALRVC  
SKLT"

## ORIGIN

Query Match 71.7%; Score 1357; DB 9; Length 3402;  
Best Local Similarity 84.1%; Pred. No. 0;  
Matches 1563; Conservative 0; Mismatches 275; Indels 21; Gaps 2;  
Qy 47 CTCTCGGCTTGGCTCGGAGAGCGCCGCGCTGCGGAGCGGCTCATCAAGTCGCTTA 106  
Db 134 CCACCTCCCTCGGCGCGAGAGCTAGCCGCGCGCTGCGGAGCGGCTCATCAAGTCGCTTA 193  
Qy 107 AACCTCTGATGAAGACGAGCGGCTGAAGCGGCGGCTCATCAAGTCGCTTA 166  
Db 194 AGCCCTTAATGAAGACGAGCGGCTGAAGCGGCGGCTCATCAAGTCGCTTA 253  
Qy 167 GCTACGAGTTCCTCGGAGAGCGCTCGGCAAGGCGGCTCATCAAGTCGCTTA 226  
Db 254 GCTACGAGTTCCTCGGAGAGCGCTCGGCAAGGCGGCTCATCAAGTCGCTTA 313  
Qy 227 AGAGCTCGGCGGCTCTGGTGGCCATCAAGTCGCTCATCAAGTCGCTTA 286  
Db 314 AGAGCTCGGCGGCTCTGGTGGCCATCAAGTCGCTCATCAAGTCGCTTA 373  
Qy 287 AGAGCTCGGCGGCTCTGGTGGCCATCAAGTCGCTCATCAAGTCGCTTA 346  
Db 374 AGAGCTCGGCGGCTCTGGTGGCCATCAAGTCGCTCATCAAGTCGCTTA 433  
Qy 347 TCATTGCGCATCCATGAAGTCTTTGAGAAATAGCAGCAAGATTGTGATTGTGATGATG 406  
Db 434 TCATTGCGCATCCATGAAGTCTTTGAGAAATAGCAGCAAGATTGTGATTGTGATGATG 493  
Qy 407 CCAGCGGAGCGGCTCTGTATGATTATCATCAAGTCGCGGCGGCTCATCAAGTCGCTTA 466

494	Db	CCAGCCGGGGCGACCTTTTATGACTACATCAGCGAGCGCAGCAGCTCAGTGAAGCGCGAAG	553
467	Qy	CCAGGCATTTCTTCCGACAGATCGTGTCTGCCCTGACACTACTGCCACACGAAACGGGATCG	526
554	Db	CTAGGCATTTCTTCCGCGAGATCGTCTCTGCCGTGCACTATTGCCATCAGACAGAGTTG	613
527	Qy	TTCAACGAGATCTCAAGCTGGAAACATCTTTCTAGATGCCAATGGAAACATCAAGATTG	586
614	Db	TCCACCAGATCTCAAGCTGGAAACATCTCTTGGATGCCAATGGGAATATCAAGATTG	673
587	Qy	CTGACTTTGGCTCTCCAACTGTACACAAAGCAAGTTCTCCACAGACTTCTGTGGGA	646
674	Db	CTGACTTTGGCTCTCCAACTGTACACCAAGCAAGTTCTGTGACACATCTGTGGGA	733
647	Qy	GCCCTCTCTACGCTCCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCCAGAGTGG	706
734	Db	GCCCTCTCTATGCTCCGCAAGATTGTCAATGGGAAGCCCTACAGGCCCCAGAGTGG	793
707	Qy	ACAGCTGGTCTCTGGGGTTCTCCTGTACATCTCTGTGATGGCACCATGCCCTTTGACG	766
794	Db	ACAGCTGGTCCCTGGGTGTTCTCCTCTACATCCTGTGTGCATGGCACCATGCCCTTTGATG	853
767	Qy	GGCAGGATCATAAACACTGGTGAAGCAATCAGTAAACGGGGCTTACGTGAGCGGCCCA	826
854	Db	GGCATGACCAATAGATCTTAGTGAACAGATCAGCAACGGGGGCTACCGGGAGCCACTA	913
827	Qy	AGCCGTCGATCGCTGTGGCTGTATCGGTGGCTGTTTAATGTGTGAACCCACACCGTCGGG	886
914	Db	AACCTCTGATCGCTGTGGCTGTATCGGTGGCTGTTTGTGTGTGAACCCACCGCCGGG	973
887	Qy	CCACACTGGAGGATGTAGCCAGTCAATTTGGTGGGTCAACTGGGGTTACAACCCGGAGTCG	946
974	Db	CCACCTGGAGGATGTGCCAGTCACTGTGTGGGTCACTGGGGCTACGCCACCCGAGTGG	1033
947	Qy	GGGAACAGGAAGCCCTCGGTAGGGTGGGCACTTGTGTGTGACTTTTGGCCGGGCTTCCA	1006
1034	Db	GAGAGCAGGAGGCTCCGATAGGGTGGGCACTCTGGCAGTGACTCTGCGCGGCTTCCA	1093
1007	Qy	TGGCGACTGGTTACGTTCGCTCTCTCGCGCCCTCTCTGGAGATGGAGCAAGGTGTGCA	1066
1094	Db	TGGCTGACTGGCTCCGGCGTTCTCTCCGCGCCCTCTCTGGAGAATGGGGCCAAAGTGTGCA	1153
1067	Qy	GCTTTCTTCAAGCAGCAGCTGTCGGGGAGGTGGAAAGCACTGTACTTGGGCTCGAGCGCAAC	1126
1154	Db	GCTTCTTCAAGCAGCATGCACTGTGTGGGGAGCACCCCTTGGCTTGGAGGGCCAGC	1213
1127	Qy	ATTCTCTTAAGAGTCCCGAAAGAGAAATGACATGTGCTCAAAATCTCAAGGTGACCCGG	1186
1214	Db	ATTCTGCTCAAGAAGTCCCGAAGAGAAATGACATGGCCAGTCTCTCCAAGTGAACCG	1273
1187	Qy	CTGAGGATACCTTTCTCGCCCTGGCAAGAGCAGCTTAAGCTTCCGAAGGCAATCTCA	1246
1274	Db	CTGATGACATGCCCCATTCGCCCCCTGGCAAGAGCAACCTCAAGCTGCCAAAGGGCATTTCTCA	1333
1247	Qy	AGAAAAGTCTCTTACCTCTGTCAGGGAGGTACAGGAGGACCTTCAGGAACTCAGACCGG	1306
1334	Db	AGAAAGGTGTGAGCTCTTCGAAAGGGGTACAGGAGGACCTTCAGGAGCTCAGCCCAA	1393
1307	Qy	TGCCTGATATCCAGGGCAGCTGTCCCTGCTGTATCCCTGCTCTCCCAAGGAAAGGCATCC	1366
1394	Db	TCCTCGAGGCCAGGGCAGGCTGCC-----CCCTGCTCCCAAGAGGGCATTC	1444
1367	Qy	TTAAGAGTCTGCACAGCGTGAATCTGGTTACTTCTCTCCAGAGCCAGCGAGTCTG	1426
1445	Db	TCAAGAAGCCCGACAGCGGAGTCTGGCTACTACTCTCTCCCGAGCCAGGTGAATCTG	1504
1427	Qy	GGAACTCTTAGACGCCAGTGATGTGTTGTGAGTGGGACCCCGTGGAGCAGAAGTCTC	1486
1505	Db	GGAGTCTTTGACGACAGGCGACGTGTTGTGAGTGGGGATCCCAAGGACGAGAGCCTC	1564
1487	Qy	CACAGGCTTCAGGGCTCCTCTCCACCGCAAGGGCATTTCTCAAACTCAATGGCAAGTTCT	1546
1565	Db	CGCAAGCTTCAGGGCTCTCTCATCGAAAGGCATCTCTCAAACTCAATGGCAAGTTCT	1624

Qy	1547	CCGCGACAGCCTTAGAAGGCACTACCCCTAGCAACCTTTGGCTCCTCTGGACAACAATGCGCT	1606
Db	1625	CCCAGACAGCCTTGGAGCTCGCGCCCCACCACTTCGCGTCCCTGGATGAATCGCCC	1684
Qy	1607	CCTCCCATCTGCAGCCCGCGCCAGCCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCC	1666
Db	1685	CACCTCGCCCCCTGGCCCGCGCCAGCCGACCTCAGGGGCTGTGAGCGAGGACAGCATCC	1744
Qy	1667	TGTCCTCCGAGTCCTTTGACCAATTGACATTGCTGAACGTCCTCCCGAAACCCCACTGA	1726
Db	1745	TGTCCTCTGAGTCCTTTGACGAGTGAGACTTGCTTGAACTGCCAGAGCCCCCACTGC	1804
Qy	1727	GGGCTGTGTCTGTGTGGACAACTCAGGGGGCTTGACAGCCTCCCTCAGAAG-----	1780
Db	1805	GGGCTGTGTGTGTGTGACAACTCAGGGGCTTGGAGAGCCCCCTCAGAGGGCCCTG	1864
Qy	1781	-----GTCTGAAGCGATGTTGGCAGGAATCTTTGGGGGATAGTGTCTTTCTGTGACAG	1834
Db	1865	GAAGCTCCCTGAGCGCTGGCGCAGGATCTTTTGGGGACAGTCTTTTCCCTGTGACAG	1924
Qy	1835	ACTGCCAAGAGGTGACTGACGCCCTACAGCAAGCCCTAGGAATCTGCTAAAGCTCAGC	1893
Db	1925	ACTGCCAAGAGGTGACACGCACTTACCGACAGGCACTGAGGGTCTGTCTAAAGCTCACC	1983

Search completed: January 25, 2005, 01:53:22  
Job time : 7951.68 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 16:12:33 ; Search time 353.728 Seconds  
(without alignments)  
9096.106 Million cell updates/sec

Title: US-09-980-464-4\_COPY\_123\_2015

Perfect score: 1893

Sequence: 1 atggagtcggtggccttact.....gaatctgctcaagctcagc 1893

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1865986 seqs, 849855171 residues

Total number of hits satisfying chosen parameters: 3731972

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA\_New.\*

- 1: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pna/US11\_NEW\_COMB.seq.\*
- 8: /cgn2\_6/ptodata/2/pna/US12\_NEW\_COMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	493.6	26.1	4933	1 PCT-US04-14421-36	Sequence 36, Appl
2	493.6	26.1	4933	7 US-10-840-512-36	Sequence 36, Appl
3	400.2	21.1	477	7 US-10-948-737-10961	Sequence 10961, A
4	221	11.7	944	7 US-10-487-901-6208	Sequence 6208, Ap
5	197.4	10.4	4094	1 PCT-US04-14421-37	Sequence 37, Appl
6	197.4	10.4	4094	7 US-10-840-512-37	Sequence 37, Appl
7	191.4	10.1	2107	7 US-10-998-525-5	Sequence 5, Appl
8	186	9.8	1512	9 US-60-638-099-231	Sequence 231, App
9	180.2	9.5	1518	9 US-60-638-099-214	Sequence 214, App
10	179.6	9.5	2006	7 US-10-998-525-21	Sequence 21, Appl
11	179.4	9.5	3394	7 US-10-995-561-364	Sequence 364, App
12	179.4	9.5	3442	7 US-10-995-561-361	Sequence 361, App
13	179.4	9.5	3466	7 US-10-995-561-363	Sequence 363, App
14	179.4	9.5	3487	7 US-10-995-561-359	Sequence 359, App
15	179.4	9.5	3514	7 US-10-995-561-350	Sequence 350, App
16	179	9.5	2123	7 US-10-998-525-11	Sequence 11, Appl
17	176.4	9.3	2179	7 US-10-956-157-4597	Sequence 4597, Ap
18	176.4	9.3	4919	7 US-10-956-157-4596	Sequence 4596, Ap
19	176.4	9.3	5074	7 US-10-990-328-3866	Sequence 3866, Ap
20	171.2	9.0	1735	1 PCT-US04-14421-15	Sequence 15, Appl
21	171.2	9.0	1735	7 US-10-840-512-15	Sequence 15, Appl
22	166.2	8.8	1527	9 US-60-638-099-106	Sequence 106, App
23	166.2	8.8	1530	9 US-60-638-099-232	Sequence 232, App
24	166.2	8.8	1948	7 US-10-998-525-3	Sequence 3, Appl

#### ALIGNMENTS

RESULT 1

PCT-US04-14421-36

; Sequence 36, Application PC/TUS0414421

; GENERAL INFORMATION:

; APPLICANT: CAENPEEL, SEAN

; APPLICANT: MANNING, GERARD

; APPLICANT: CHARDICZAK, GLEN

; APPLICANT: HIGORIEV, IGOR

; TITLE OF INVENTION: NOVEL KINASES

; FILE REFERENCE: 034536-1454

; CURRENT APPLICATION NUMBER: PCT/US04/14421

; CURRENT FILING DATE: 2004-05-07

; PRIOR APPLICATION NUMBER: 60/469,014

; PRIOR FILING DATE: 2003-05-09

; NUMBER OF SEQ ID NOS: 239

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 36

; LENGTH: 4933

; TYPE: DNA

; ORGANISM: Mus musculus

; PCT-US04-14421-36

Query Match 26.1%; Score 493.6; DB 1; Length 4933;

Best Local Similarity 66.6%; Pred.No. 1.1e-97;

Matches 766; Conservative 0; Mismatches 354; Indels 30; Gaps 3;

Qy	71	CCCGCGCGCTGGCGGACGGGCTCATCAAGTCGCTAAACCTCTGATGAAGAAGCAGCGG	130
Db	68	CGCTGGAGCGGTGGCGGGCGACTGTCAGCCAGTGGAGCCAGGAACCGCACGGG	127
Qy	131	TGAAGCGGACCATCAACAAACAACTGGCGGACCGCTACGAGTTCTTGGAGACGCTGG	190
Db	128	TGAAGCGGACATCAACCAACAACTTGAAGCATCGCTATGAGCTGCAAGAGACCGCTGG	187
Qy	191	GCAAGGCGACCTACGGGAAGGTGAAGAAGCCAGAGAGAG--CTCGGGCGCTCTGGTG	247
Db	188	GCAAGGTACTCTAGGCAAGTCAAGAGAGCCACCGAGAGGTTTTCAGGCCGAGTGGTTG	247
Qy	248	CCATCAAGTCCATCAGGAAGACAAATCAAGATGAGCAGGATCTGCTGCACATACGGA	307
Db	248	CTATAAATCCATCCGTAAGGACAAATTAAGCAGGCTAGACATGTTTCACATCAGAC	307
Qy	308	GGAGATTTGAGATCATGTCTTCTACTCAACACCCCAACATCATTCATCATCAAGTGT	367
Db	308	GAGAGATTGAGATCATGTCTCTCCCTCAACCATCTCATATCATCATCAGTATTTATGAAGTGT	367







Db 119 CGGCAAACTCTGGGCAATGGATCTTTGGCAAGTGAAGATCGCGGAGCATATCTAC 178  
Qy 234 GGGGCTCTGGTGGCCATCAAGTCATCAGGAAGACAAATCAAGATGAGCGGATCT 293  
Db 179 CGGCGACAAGGTGCGGATCAAGATCTCTCAACCGCAGGAAGATCAAGCCATGGAATGGA 238  
Qy 294 GCTGCACATACGGAGGAGATGAGATCATGCTCTCACTCAACACCCACATCATTCG 353  
Db 239 AGAGAAAGTTCGAGAGAGATTAATCTGAGGTGTTTATGATCCACACATCATTAAG 298  
Qy 354 CATCATGAAGTGTGAGATAGCAGCAAGATTTGATTTGATGAGTATGCGAGCG 413  
Db 299 GTTGTATGAAGTGTGAGAGACGACACAGACATTTACGTGTCATGAGTACGCAAGTC 358  
Qy 414 AGGCGATCTGATGATACATCAGTACGCGCCACGCTGAGTACGCGGAGCGCAGGCA 473  
Db 359 GGGAGAGCTCTTCGACTACATTTGTTGAGATGGTTCGCTTTCATGAAGATCAAGCTCGACG 418  
Qy 474 TTTCTTCGACAGATCGTGTCTGCGCTGCATCTACTGCCACACGAGGATCGTTCAACG 533  
Db 419 TTTCTTCGACAGATATATCAGGGTCGAGTACTGCCATAGAAACATGTTAGTTCATCG 478  
Qy 534 AGATCTCAAGCTGGAAACATCTCTTCTAGATGCCAATGGAAACATCAAGATTTGCTGACTT 593  
Db 479 TGATCTCAAGCGGAAATCTGCTCTCGACTCTCAATGGAACGTTAAAGATTGCGGACTT 538  
Qy 594 TGGCTCTCCAACTGTACCAAAAGCAAGTCTCTCAGACGTTCTGTGGGAGCCCTCT 653  
Db 539 TGGCTTAAAGCAACATTTATGCGAGACGACATTTCTTAAACATAGCTGTGGAAGTCCAAA 598  
Qy 654 CTACGCTCGCTCGAGATAGTCAACGGGAAGCCCTATGTGGGCCACAGAGTGGACAGCTG 713  
Db 599 CTATGCTCTCAGAGGTTATCTCTGGCAAGCTCTATGCTGGAACAGAGTTGACGTATG 658  
Qy 714 GTCTCTGGGCTTCTCTGTACATCTCTGTGTCATGTCATGACACCTTTGTCGGGAGGA 773  
Db 659 GAGTTGGGCGTTATTTGTATGCACTTCTTTGGCGAGCCTTCCATTTGACGATGAAA 718  
Qy 774 TCATAAACACTGCTGAGCAATCAGTAACGGGGCTTACCGTACGCGGCCAAGC---C 830  
Db 719 CATTCAAAACCTTTTCAAGAAATTAAGGGTGGCATCTACATCTACAAAGTCTCTC 778  
Qy 831 GTCCGATCGCTGTGGCCTGATCCGCTGCTGTTAATGTTGAACCCCAACCGTGGGCGAC 890  
Db 779 TGCCGGTGCAAAAGATTTAATCCCGCTATGCTAGTGGTTGATCCATGAACGATGAC 838  
Qy 891 ACTGGAGATGATAGCCAGTCAATGGTGGTCAA 923  
Db 839 TATAGCGAGATTCGTGAGCATCCTTGGTTCCA 871

## RESULT 5

PCT-US04-14421-37  
; Sequence 37, Application PC/TUS0414421  
; GENERAL INFORMATION:  
; APPLICANT: SUGEN, INC.  
; APPLICANT: CAENEPEEL, SEAN  
; APPLICANT: MANNING, GERARD  
; APPLICANT: CHARYDCZAK, GLEN  
; APPLICANT: GRIGORIEV, IGOR  
; TITLE OF INVENTION: NOVEL KINASES  
; FILE REFERENCE: 034536-1454  
; CURRENT APPLICATION NUMBER: PCT/US04/14421  
; PRIOR FILING DATE: 2004-05-07  
; PRIOR APPLICATION NUMBER: 60/469,014  
; NUMBER OF SEQ ID NOS: 239  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 37  
; LENGTH: 4094  
; TYPE: DNA  
; ORGANISM: Mus musculus

## PCT-US04-14421-37

Query Match 10.4%; Score 197.4; DB 1; Length 4094;  
Best Local Similarity 54.7%; Pred. No. 1.9e-33;  
Matches 415; Conservative 0; Mismatches 341; Indels 3; Gaps 1;  
Qy 168 CTAGGAGTCTCTGGAGACGCTGGCGAGGCGACCTACGGGAAGTGAAGAGCACGAGA 227  
Db 158 CTAGGAGATCGACGCGACCAATCGCGAAGGCAATCTTCGCTGTGTCGCGGCGCACGCA 217  
Qy 228 GAGTCGCGGGCGTCTGCTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATAGCA 287  
Db 218 CCTCGTCACCAAGCGCAAGGTTCCTATTAATCATAGATAAGCCAGCTGGATGAAGA 277  
Qy 288 GGATCTCTGCACATACGAGGAGGAGATGAGATCATGTCTTCACTCAACACCCCAACAT 347  
Db 278 AAACTTGAAGAAGATTTTCGCGGAGGTTTCAGATAATGAAGATGCTTTGCCATCCACAT 337  
Qy 348 CATTGCCATCCATCAAGTGTTCAGAAATAGCAGCAAGATTTGATTTGTCTATGGAGTATGC 407  
Db 338 CATCAGACTCTACAGGTCTAGGAGACAGAGCGATGATTTACTTGGTGACAGATACGC 397  
Qy 408 CAGCGAGGCGATCTGTATGATTAATCATCAGTACGCGGCCACCGCTGAGTACGCGGAGCG 467  
Db 398 TAGCGGAGGCGAGATTTTGACCACTTGGTAGCCCATGGAAGAAATGCGAGAGAAGGAGC 457  
Qy 468 CAGGCATTTCTTCGACAGATCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527  
Db 458 TCAGCGGAAGTTCAAAACAGATCGTCAACGCGGTGTATTTTGTCTGCTGCGAATATCGT 517  
Qy 528 TCACCGAGATCTCAAGCTGGAAACATCTTCTTAGATGCCAATGGAACATCAAGATTGC 587  
Db 518 TCATCGTATTTAAAGCGGAAACCTTACTTCTGGAGCGCAATCTGAAATATCAAAATAGC 577  
Qy 588 TGACTTTGGCTCTCCAACCTGTACACAAAGCAAGTTCTCTCCAGACGTTCTGTGGGAG 647  
Db 578 AGACTTTGGCTTCAGCAACCTCTTCACTCCAGGCGAGCTGCTGAAGACGTTGGTGGCAG 637  
Qy 648 CCCTCTCTAGCGCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCAAGSTGGA 707  
Db 638 CCCTCTCTATGCGCGCCACAGCTCTTCAAGAAAGGAAGAAATATGATGGGCGCCAAAGTGA 697  
Qy 708 CAGCTGCTCTCTGGGCTTCTCTCTGATACATCTCTGTCATGTCGACCATGCTTTCACGG 767  
Db 698 CATATGAGCGCTTGGAGTTGCTCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 757  
Qy 768 GCAGGATCATAAACACTGCTGTAAGCAATCAGTAACGGGGCTTACCGTACGCGCCCAA 827  
Db 758 GAGCACACTGCAGAAATCTGCGGCGCGCTGCTGAGTGGCAAGTTCCGATCCCGTCTCT 817  
Qy 828 GCCGTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 884  
Db 818 TATGTCACAGAGTGTGAGCATTGATCCGCCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 877  
Qy 885 GGCACACTGGAGATGTAGCCAGTCAATTTGGTGGGTCAA 923  
Db 878 CCTCTCAATGAACACATCTGACGAGCACAAGTGGATGAA 916

## RESULT 6

US-10-840-512-37  
; Sequence 37, Application US/10840512  
; GENERAL INFORMATION:  
; APPLICANT: CAENEPEEL, SEAN  
; APPLICANT: MANNING, GERARD  
; APPLICANT: CHARYDCZAK, GLEN  
; APPLICANT: GRIGORIEV, IGOR  
; TITLE OF INVENTION: NOVEL KINASES  
; FILE REFERENCE: 034536-1455  
; CURRENT APPLICATION NUMBER: US/10/840,512  
; PRIOR FILING DATE: 2004-05-07  
; PRIOR APPLICATION NUMBER: 60/469,014  
; PRIOR FILING DATE: 2003-05-09



QY 731 TGTATCTCTGGTGCATGCCACCATGCTTTGAGCGGCGAGATCATAAACACTGGTGA 790  
 Db |||||  
 QY 834 TTTATGCTCTTTGTTGGCACTCTCCCATTTGAGATGAGATATTCAAACCTTTTCA 893  
 Db |||||  
 QY 791 AGCAAAATCA 799  
 Db |||||  
 QY 894 AGAAAATAA 902

## RESULT 8

US-60-638-099-231  
 ; Sequence 231, Application US/60638099  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgerton, Michael D  
 ; TITLE OF INVENTION: Transgenic Plants with Enhanced Agronomic Phenotypes  
 ; FILE REFERENCE: 38-21(53720)  
 ; CURRENT APPLICATION NUMBER: US/60/638,099  
 ; CURRENT FILING DATE: 2004-12-21  
 ; NUMBER OF SEQ ID NOS: 48056  
 ; SEQ ID NO 231  
 ; LENGTH: 1512  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 US-60-638-099-231

Query Match 9.8%; Score 186; DB 9; Length 1512;  
 Best Local Similarity 56.5%; Pred. No. 4.6e-31;  
 Matches 367; Conservative 0; Mismatches 280; Indels 3; Gaps 1;

QY 153 CAACCTCGGGCACCGCTACGAGTTCTTGGAGACGCTGGGCAAGGCGACCTACGGGAAGGT 212  
 Db |||||  
 QY 24 CAACCTTTGGGCGTTTACCGAATTTGCCAAACCTTAGGAATTTGGTTCGTTNAAGT 83  
 Db |||||  
 QY 213 GAAGAACGCGAGAGAGCT---CGGGCGTCTGGTGCCCATCAAGTCCATCAGGAAGA 269  
 Db |||||  
 QY 84 GAAGATCGCGAACATATATTGACTGGCCATAAGGTGGCAATCAAGATCTCTCAATCGCCG 143  
 QY 270 CAAAATCAAGATGAGCAGGATCTGTCACATACGAGGAGAGATTCAGATCATGTCTTC 329  
 Db |||||  
 QY 144 TAAGTCAAGAGCATGGAATGGAAGAGAAAGTGAAGAGAAATCAAGTACTTGAGATT 203  
 QY 330 ACTCAACACCCCAATCATTTGCCATTCATGAAGTGTGTTGAGAAATAGCAGCAAGATTGT 389  
 Db |||||  
 QY 204 ATTTATGATCTCTCATATACATACGCTTTATGAGTGTATAGATACACCTGCTGATTTA 263  
 QY 390 GATTGTCATGAGTATGCCAGCGAGGATCTGTATGATTAATCATAGTGAGCGGCCACG 449  
 Db |||||  
 QY 264 TGTGTTATGGAGTATGTTAAATCTGGAGAGTTGTTTGAATACATTTGTTGAGAGGGAAG 323  
 QY 450 GCTGAGTGAGCGGACGCCAGGCAATTTCTTCCGACAGATCGTGTCTGCCCTGCACTACTG 509  
 Db |||||  
 QY 324 ACTACATGAAGAGAGCGCGCGCTTTTTCAGCAGATCATATCTGTTGTTGATATTG 383  
 QY 510 CCACAGAACGGGATCGTTTCAACGAGATCTCAAGCTGGAAACATCTCTTAGATGCCAA 569  
 Db |||||  
 QY 384 CCATAGGAACATGTTGTCACCGTGATTTAAAGCCAGAGAAATCTCTTTTGGATTGAA 443  
 QY 570 TGGAAACATCAAGATTGCTGACTTTGGCTCTCAACCTGTACCACAAAGCAAGTTCTCT 629  
 Db |||||  
 QY 444 ATGCAACATTAAGATTTGCTGATTTTGGCTTAAGTAATGTTATGCGTGTATGCTCACTTCT 503  
 QY 630 CCAGACCTTTCTGTTGGAGCGCTCTCTACGCTCGCTCGATAGTCAACGGGAAGCCCTA 689  
 Db |||||  
 QY 504 TAAGACGAGTTGGTAGCCGATTTATGAGCAGCTGAGGTCAATCTGTTAACTATA 563  
 QY 690 TGTGGGCGCGAGGTGAGCAGTGGTCTCTGGGGTCTCTCTGTACATCTCTGTGTCATGG 749  
 Db |||||  
 QY 564 TGTGTTCTCTGAAGTTGATGTTGGAGCTGGCGGTTTATCTTTATGCTCTCTCTTTGTTGG 623  
 QY 750 CACCATGCCCTTTGACCGGAGGATCATAAACACTGTTGAAGCAATCA 799  
 Db |||||  
 QY 624 CACTCTCCATTTGACCATGAGAAATATCTTAAACCTTTTAAAGAAAATAA 673

## RESULT 9

US-60-638-099-214  
 ; Sequence 214, Application US/60638099  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgerton, Michael D  
 ; TITLE OF INVENTION: Transgenic Plants with Enhanced Agronomic Phenotypes  
 ; FILE REFERENCE: 38-21(53720)  
 ; CURRENT APPLICATION NUMBER: US/60/638,099  
 ; CURRENT FILING DATE: 2004-12-21  
 ; NUMBER OF SEQ ID NOS: 48056  
 ; SEQ ID NO 214  
 ; LENGTH: 1518  
 ; TYPE: DNA  
 ; ORGANISM: Oryza sativa  
 US-60-638-099-214

Query Match 9.5%; Score 180.2; DB 9; Length 1518;  
 Best Local Similarity 55.9%; Pred. No. 8.4e-30;  
 Matches 363; Conservative 0; Mismatches 283; Indels 3; Gaps 1;

QY 154 AACCTCGGCGACCGCTACGAGTTCTTGGAGACGCTGGGCAAGGCGACCTACGGGAAGGTG 213  
 Db |||||  
 QY 25 AACCTCTTGGCGTTACCGGATTGGCAAAACCTTAGGATTGGTCAATTTGGCAAGTG 84  
 Db |||||  
 QY 214 AAGAACGCGAGAGAGCT---CGGGCGTCTGGTGCCCATCAAGTCCATCAGGAAGAC 270  
 Db |||||  
 QY 85 AAGATCGCGGAGCATATATTGACTGGTCAAGGTGGCAATCAAGATCTCTCAATCGCCGT 144  
 QY 271 AAAATCAAGATGAGCAGGATCTGTCACATACGAGGAGAGATTCAGATCATGTCTTCA 330  
 Db |||||  
 QY 145 AAGTCAAGAGCATGAGATGGAGAGAAAGTTAAAGAGAAATCAAGTACTTAGATT 204  
 QY 331 CTCAACACCCCAATCATTTGCCATTCATGAAGTGTGTTGAGAAATAGCAGCAAGATTGTG 390  
 Db |||||  
 QY 205 TTTATGACCCCAATCATTTGCTTTATGAGTGTAGTATGACACCCAGCTGATATTAT 264  
 QY 391 ATTGTCTAGGATGATGCCAGCGGAGTCTGTATGATTACATCAGTGAGCGGCCACGG 450  
 Db |||||  
 QY 265 GTTGTATTGGAGTATGTCAAACTGGAGAGTTGTTTGAATACATCGTTGAGAGGGAAGA 324  
 QY 451 CTGAGTGAGCGGACGCCAGGCAATTTCTTCCGACAGATCGTGTCTGCCCTGCACTACTGC 510  
 Db |||||  
 QY 325 CTGCAAGAGAGAGAACTCGACGCTTTTCCAGCAGATCATATCTGTTGTTGATATTGC 384  
 QY 511 CACAGAACGGGATCGTTTCAACGAGATCTCAAGCTGGAAACATCTCTTAGATGCCAAT 570  
 Db |||||  
 QY 385 CATAGAAACATGTTGTTGTTTCTGATCTTAAGCCAGAGAACCTTCTTTTGGACTCCAA 444  
 QY 571 GGAACATCAAGATTGCTGACTTTGGCTCTCCAACTGTACCACAAAGCAAGTTCTTC 630  
 Db |||||  
 QY 445 TGAATGTTAAGATTGACAGCTTTGGCTTGAAGTAATGTTATGCGTGTATGCTCTTCTG 504  
 QY 631 CAGACGCTTCTGTTGGAGCGCTCTCTACGCTCGCTCGATAGTCAACGGGAAGCCCTAT 690  
 Db |||||  
 QY 505 AAGACAAGTTGTTAGTACCCAAATTTATGAGCAGCTGAGGTGATATCTGTTAACTATAT 564  
 QY 691 GTGGGCGCGAGGTGAGCAGCTGCTCTGCGGGTCTCTGTTGAGATCTCTGTTGAGTGGC 750  
 Db |||||  
 QY 565 GCTGGCGCTGAAAGTTGATGTTGTTGAGTGTGTTGTTATCTTTATGCTCTCTTCTGTT 624  
 QY 751 ACCATGCCCTTTGACCGGAGGATCATAAACACTGTTGAAGCAATCA 799  
 Db |||||  
 QY 625 ACCCTTCCATTTGATGACGAGAAATATCCCAACCTTTTAAAGAAAATAA 673

## RESULT 10

US-10-998-525-21  
 ; Sequence 21, Application US/10998525  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Allen, Stephen  
 ; APPLICANT: Heppard, Elmer  
 ; APPLICANT: Sakai, Hajime

```
; APPLICANT: Weng, Zude
; APPLICANT: Helentjaris, Tim
; APPLICANT: Macool, Daniel
; APPLICANT: Miao, Guo-Hua
; TITLE OF INVENTION: Plant Carbon Catabolite Repression Proteins
; FILE REFERENCE: BB-1315
; CURRENT APPLICATION NUMBER: US/10/998,525
; CURRENT FILING DATE: 2004-11-29
; PRIOR FILING DATE: US/09/857,522
; PRIOR FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: PCT/US99/29824
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: 60/112,563
; PRIOR FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2006
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-998-525-21

Query Match          9.5%; Score 179.6; DB 7; Length 2006;
Best Local Similarity 55.8%; Pred. No. 1.2e-29;
Matches 363; Conservative 0; Mismatches 284; Indels 3; Gaps 1;

Qy 153 CAACCTCGCGCACCCTACGAGTTCCTGGAGAGCTGGGCAAGGGACCTACGGGAGGT 212
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 213 GAAGAAGGC---ACGAGAGAGCTGGGGCGCTCTGGTGGCCATCAAGTCCATCAGGAAGA 269
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 146 CAAGATTGCGGAGCATATAAAACTGTGTCAAGGTGGCGGTCAAGATCTTAAACGCGC 205
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 270 CAAAATCAAGATGAGCAGAGATCTGTGCATACGAGGAGGATGAGATCATGTCTTC 329
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 206 GAAATCAAAAACATGGAGATGGAAGAGAAAGTGAAGAGAGATCAAGATATTAAGATT 265
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 330 ACTCAACACCCACATCATTCATCCATGAAGTCTTTGAGAAATAGCAGCAAGATTGT 389
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 266 ATTATGACCCACATATCATCCGCTTTAAGAGTATAGAGCCACGACTGATATTA 325
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 390 GATTGTGATGAGTATGCCAGCGGCGATCTGTATGATTACATCAGTGAAGCGGCACG 449
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 326 TGTGTTATGAGTATCTTAAGTCTGTTGAATGTTTGAATACATCTTGCAGAAAGTAG 385
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 450 GCTGAGTGAAGGAGCGCAGGATTTCTTCGAGAGATCGTGTCTGCCCTGCACTACTG 509
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 386 GCTACAGGAGGAGAGGCGCGCGTTCTTTCAACAGATCATATCTGGTGTCAATATTG 445
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 510 CCACGAAACGGGATCTTCACCGAGATCTCAAGCTCGAAGAAACATCTCTAGATGCCAA 569
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 446 CCACAGGAAATGTTGTTGACCGCGATCTAAAGCGGGAACCTCTTTTGGACAATAA 505
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 570 TGAAGAACATCAAGATTCTGACTTTGGCCCTCTCAACCTGTACCAAAAGGCAAGTTCTCT 629
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 506 TTGTGATTTAAGATTGCGGATTTTGGCTTAAGTATGTTATGCGTGAAGCCACTTTCT 565
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 630 CCAGACGTTCTGTGGAGCCCTCTCTACGCTCGCTGAGATAGTCAACGGGAAGCCCTA 689
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 566 TAAGACAAGTTGTGTAGCCCAATATATGAGCTCCGAGGTTATATCTGGAAGAACTGTA 625
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 690 TGTGGGCGCAGAGTGACAGCTGTTCTGTGGGGTCTCTGTACATCTGTGTGATGG 749
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 626 CGTGGGCGCTGAAGTTGATGTATGAGAGCTCGGTTGTTATTTATGCTCTTCTATGTTG 685
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 750 CACCATGCTTTTGACGGGAGGATCATAAAACACTGGTGAAGCAAAATCA 799
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 686 TACTCTTCATTGATGATGAGAACATACCAACCTTTTAAAGAAATAA 735
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-10-995-561-364
```

```
; Sequence 364, Application US/10995561
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 364
; LENGTH: 3394
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-364

Query Match          9.5%; Score 179.4; DB 7; Length 3394;
Best Local Similarity 53.1%; Pred. No. 1.5e-29;
Matches 403; Conservative 1; Mismatches 352; Indels 3; Gaps 1;

Qy 168 CTACGAGTTCTCGGAGACGCTGGCAAGGCGACCTACGGGAAGGTGAAGAGCAGAGA 227
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 813 CTACAGACTGTTGAAAACAATCGCAAGGGGAATTTTGAAAAAGTAAATTTGCAAGACA 872
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 228 GAGCTCGGGCGCTCTGGTGGCCATCAAGTCCATCAGGAAAGACAAAATCAAGATGAGCA 287
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 873 TATCCTTTACAGGAGAGAGGTTGCAATAAAATAATTGACAAAATCACTAGTTGAATCCAAC 932
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 288 GGAATCTGTGCACATACGAGGAGATTTGAGATCATGTCTTCTCACTCAACACCCCCACAT 347
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 933 AAGTCTACAAAAGCTCTTCAGAGAGTAAGAATAATGAAGATTTTAAATCATCCCAATAT 992
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 348 CATTCGCATCCATGAAGTGTTCAGATAGCAGCAAGATGTGATTTCTCATGAGTATGC 407
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 993 AGTGAAGTTATTTCGARGTCATTGAAACTGAAAAACACTTCACTAATCATGGAATATGC 1052
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 408 CAGCCGAGGCGATCTGTATGATTACATCATGAGCGGCCACGCTGAGTGAAGGGAACGC 467
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1053 AAGTGAAGGTGAAGTATTGACTATTGTTGGTTCACATGCGAGATGAAGGAAAAGAGC 1112
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 468 CAGGCATTTCTCCGACAGATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 527
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1113 AAGATCTAAATTTAGACAGATTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1172
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 528 TCACCGAGATCTCAAGCTGGAAGAACATCTTCTAGATGCCAATGGAACATCAAGATTGC 587
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1173 ACATCGAGACCTCAAGGCTGAAATCTATTGTTAGATGCCGATATGAACATTAATAATAG 1232
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 588 TGACTTTGGCTCTCCAACTGTACCAAAAGGCAAGTTCTCTCCAGACGTTCTGTGGGAG 647
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1233 AGATTTGCGTTTACGCAATGAATTTACTGTTGGCGGTAACTCGACACGTTTGTGGCAG 1292
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 648 CCCTCTCTACGCTCGCTGAGATAGTCAACGGGAAGCCCTATATGTGGGCCAGAGGTGA 707
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1293 TCCTCCATACGACGACCTGAGCTCTTCAGGGCAAGAAATAATGACGGGCGAGAAGTGA 1352
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 708 CAGCTGGTCTCTGGGGTCTCTCTGATCATCTGTTGATGCGACCATGCTGCTTTGACGG 767
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1353 TGTGTGGAGTCTGGGGTCTATTTATACACACTAGTCAGTGGGCTCACTTCCCTTTGATGG 1412
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 768 GCAGGATCATAAAACATGTTGAAGCAAAATCAGTAACGGGCTTACCGTGAGCGCCCAA 827
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1413 GCAAAACCTTAAGAACTGAGAGAGAGATATTAAGAGGGAATATAGAAATTCCTTCTA 1472
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 828 GCGTCCGATGCTGTG---GCTGATCCGGTGGCTGTTAAATGGTGAACCCACCCGTCG 884
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1473 CATGTCTACAGACTGTGAAACCTTCTCAACCGTTTCTCGTGTCTAAATCCAATTAACG 1532
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 885 GGCACATGAGGAGATGTAGCCAGTCAATTTGGTGGTCA 923
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1533 CGGCACTCTAGAGCAAAATCATGAAGGACAGGTGGATCA 1571
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 12
US-10-995-561-361
; SEQUENCE 361, Application US/10995561
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 361
; LENGTH: 3442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-361

Query Match          9.5%; Score 179.4; DB 7; Length 3442;
Best Local Similarity 53.1%; Pred. No. 1.5e-29;
Matches 403; Conservative 1; Mismatches 352; Indels 3; Gaps 1;

QY 168 CTACGAGTCTCTGGAGACGCTGGGCAAGGCGCACCTACCGGAAGGTGAAGAGGCACGAGA 227
DB 813 CTACAGACTCTTGAACAACATCGCAAGGGGAATTTGCAAAAGTAAATTTGGCAGACA 872
QY 228 GAGTCGGGGCGTCTGGTGGCCATCAAGTCCATCAGGAAGACAAATCAAAAGATGACA 287
DB 873 TATCTCTTACAGGACAGAGGTTGCAATAAAATAATTTGCAAAAGTAAATTTGGCAGACA 872
QY 288 GATCTGCTGCACATACGAGGAGGAGATTGAGATCATGTCTTCACTCAACCCACCCACAT 347
DB 933 AAGTCTACAAAGCTCTTCAGAGAAGTAAAGATTAATGAAGATTTTAAATCATCCCAATAT 992
QY 348 CATTCGCCATCCATGAAGTGTGGTGAAGTACGAGCAAGATTTGATTTGATGAGATGATGC 407
DB 1053 AAGTGGAGGTGAAGTATTTGACTATTTGGTTGCACATGGCAGGATGAAGGAAAAAGAGC 1112
QY 468 CAGCCGAGCGCATCTGTATGATTACATCAGTACGCGGCCACGCTCAGTACGAGCGGACGC 467
DB 1113 AAGATCTAAATTTAGACAGATTGTGTCGAGTTCAATCTGCTCAGTACAGAAACGGATCGT 1172
QY 528 TCACCGAGATCTCAAGCTGGAAAAACATCCTCTTAGATGCAATGGAATGGAACATCAAGATTGC 587
DB 1233 AGATTTGGGTTTACGCAATGAATTTACTGTTGGCGGTAAACTCGACACGTTTGTGGCAG 1292
QY 648 CCCTCTACGCTCCGCTCAGATAGTCAACGGGAAGCCCTATGTGGGCCACGAGGTGGA 707
DB 1293 TCCTCCATACGACGACCTGAGCTCTTCAGGCAAGAAATATGACGGGCCAGAGTGA 1352
QY 708 CAGCTGGTCTCGGGGCTTCTCTGTACATCTCTGTGTCATGCGACCATGCGCTTTCACGG 767
DB 1353 TGTGTGGAGTCTGGGGGTCAATTTATACACACTAGTACGCTAGTGGCTCACTTCCCTTCTA 1412
QY 768 GCAGGATCATAAACACTGTGAAGCAATCAGTAAACGGGCTTACCGTGAGCGGCCCAA 827
DB 1413 GCAAAACCTTAAGGAACCTGAGAGAGAGATTTAAGAGGGAATACAGAAATCCCTTCTA 1472
QY 828 GCCGTCCGATCGCTGTG---GCCGTGATCGGTGGCTGTTAATGGTGAAACCCACCGCTCG 884
DB 1473 CATGCTACAGACTGTGAAACCTTCTCAACGCTTTCCTGGTGTAAATCCAAATTAACG 1532
QY 885 GGCCACACTGGAGGATGTAGCCAGTCAATTTGGTGGGTCAA 923

RESULT 13
US-10-995-561-363
; SEQUENCE 363, Application US/10995561
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 363
; LENGTH: 3466
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-363

Query Match          9.5%; Score 179.4; DB 7; Length 3466;
Best Local Similarity 53.1%; Pred. No. 1.5e-29;
Matches 403; Conservative 1; Mismatches 352; Indels 3; Gaps 1;

QY 168 CTACGAGTCTCTGGAGACGCTGGGCAAGGCGCACCTACCGGAAGGTGAAGAGGCACGAGA 227
DB 813 CTACAGACTGTGTAACAACATCGCAAGGGGAATTTGCAAAAGTAAATTTGGCAGACA 872
QY 228 GAGTCGGGGCGTCTGGTGGCCATCAAGTCCATCAGGAAGACAAATCAAAAGATGACA 287
DB 873 TATCTCTTACAGGACAGAGGTTGCAATAAAATAATTTGCAAAAGTAAATTTGGCAGACA 932
QY 288 GATCTGCTGCACATACGAGGAGGAGATTGAGATCATGTCTTCACTCAACCCACCCACAT 347
DB 933 AAGTCTACAAAGCTCTTCAGAGAAGTAAAGATTAATGAAGATTTTAAATCATCCCAATAT 992
QY 348 CATTCGCCATCCATGAAGTGTGGTGAAGTACGAGCAAGATTTGATTTGATGAGATGATGC 407
DB 993 AGTGAAGTATTTGARGTCACTTGAACACTGAAAAACACTCTACCTAATCATGGAATATGC 1052
QY 408 CAGCCGAGCGCATCTGTATGATTACATCAGTACGCGGCCACGCTCAGTACGAGCGGACGC 467
DB 1053 AAGTGGAGGTGAAGTATTTGACTATTTGGTTGCACATGGCAGGATGAAGGAAAAAGAGC 1112
QY 468 CAGGCAATTTCTCCGACAGATCGTGTCTGCCCTGCACTACTGCCACAGAACCGGATCGT 527
DB 1113 AAGATCTAAATTTAGACAGATTGTGTCGAGTTCAATCTGCTCAGTACAGAAACGGATCGT 1172
QY 528 TCACCGAGATCTCAAGCTGGAAAAACATCCTCTTAGATGCAATGGAATGGAACATCAAGATTGC 587
DB 1173 ACATCGAGACCTCAAGCTGAAATCTATTTAGATGCGGATATGAACATTAATATAGC 1232
QY 588 TGACTTTGGCTCTCCAACTGTACCAAAAGCAAGTTCTTCCAGACGTTCTGTGGGAG 647
DB 1233 AGATTTGGGTTTACGCAATGAATTTACTGTTGGCGGTAAACTCGACACGTTTGTGGCAG 1292
QY 648 CCCTCTACGCTCCGCTCAGATAGTCAACGGGAAGCCCTATGTGGGCCACGAGGTGGA 707
DB 1293 TCCTCCATACGACGACCTGAGCTCTTCAGGCAAGAAATATGACGGGCCAGAGTGA 1352
QY 708 CAGCTGGTCTCTGGGCGTCTCTGTATACATCTCTGTGTCATGCGACCATGCGCTTTCACGG 767
DB 1353 TGTGTGGAGTCTGGGGGTCAATTTATACACACTAGTACGCTAGTGGCTCACTTCCCTTCTA 1412
QY 768 GCAGGATCATAAACACTGTGAAGCAATCAGTAAACGGGCTTACCGTGAGCGGCCCAA 827
DB 1413 GCAAAACCTTAAGGAACCTGAGAGAGAGATTTAAGAGGGAATACAGAAATCCCTTCTA 1472
QY 828 GCCGTCCGATCGCTGTG---GCCGTGATCGGTGGCTGTTAATGGTGAAACCCACCGCTCG 884
DB 1473 CATGCTACAGACTGTGAAACCTTCTCAACGCTTTCCTGGTGTAAATCCAAATTAACG 1532
QY 885 GGCCACACTGGAGGATGTAGCCAGTCAATTTGGTGGGTCAA 923
```



Db 1473 CATGCTTACAGACTGTGAAACCTTCTCAAAACGTTTCTCGTGTCTAAATCAATTAACG 1532  
QY 885 GGCCACACTGGAGGATAGCAGTCAATGTTGGTCA 923  
Db 1533 CGGCACCTTAGACCAATCATGAAGCAGGTGGATCAA 1571

## RESULT 14

US-10-995-561-359  
; Sequence 359, Application US/10995561  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; TITLE OF INVENTION: DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001559  
; CURRENT APPLICATION NUMBER: US/10/995.561  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 359  
; LENGTH: 3487  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-995-561-359

Query Match 9.5%; Score 179.4; DB 7; Length 3487;  
Best Local Similarity 53.1%; Pred. No. 1.5e-29;  
Matches 403; Conservative 1; Mismatches 352; Indels 3; Gaps 1;  
QY 168 CTACGAGTTCCTGGAGCGCTGGCAAGGCGACCTACGGGAAGGTGAAGAGGCACGAGA 227  
Db 813 CTACAGACTGTGAAACAAATCGCAAGGGGAATTTTGCAAAAGTAAATTTGCAAGACA 872  
QY 228 GAGCTCGGGCGTCTGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGCA 287  
Db 873 TATCCTTACAGGAGAGAGTTCGCAATAAATAATTTGCAAAACTCAGTTGAATCAAC 932  
QY 288 GGATCTGCTGCACATACGGAGGAGATTGAGATCATGCTTCACTCAACCCACCCACAT 347  
Db 933 AAGTCTACAAAGCTCTTCAGAGAGTAAAGATAAGAAATTTAAATCATCCCAATAT 992  
QY 348 CATTGCCATCCATGAAGTGTGAGATAGCAGCAAGATTGTGATGTCTATGAGTATGC 407  
Db 993 AGTGAAGTTATTCGARGTATTGAACTGAAAAAACACTCTACTTAATCATGAATATGC 1052  
QY 408 CAGCCGAGGCGATCTGTATGATTACATCAGTGAAGCGCCACGGCTGAGTGGGAGCG 467  
Db 1053 AAGTGAAGTGAAGTATTTGACTATTTGGTTGCACATGGCAGGATGAAGAAAAGAGC 1112  
QY 468 CAGGCAATTTCTCCGACAGATCGTCTGCCCTGCACTAGTGCACCAAGCAAGCGGATCGT 527  
Db 1113 AAGATCTAAATTTAGACAGATTGTCTGCAGTTCAATATCTGCATCAGAAACGGATCGT 1172  
QY 528 TCACCGAGATCTCAAGCTGGAAACATCTTCTAGATGCCAATGGAAACATCAAGATTGC 587  
Db 1173 ACATCGAGACTCAAGGCTGAAATCTATTGTTAGTGGCGATATGAACATTAATAATGC 1232  
QY 588 TGACTTTGGCCTCTCCACCTGTACCAAGGCAAGTTCTCCAGACGTTCTGTGGGAG 647  
Db 1233 AGATTTGGGTTTATGCAATGAAATTTACTGTTGGCGGTAAACTCGACACGTTTGTGGCAG 1292  
QY 648 CCCTCTACGCTCGCTCAGATAGTCAACCGGAAGCCCTATGTGGGCCACAGAGTGA 707  
Db 1293 TCCTCCATACGAGCACTGAGCTCTTCAGGGCAAGAAATATGACGGGCCAGAGTGA 1352  
QY 708 CAGCTGGTCTCTGGGCGTCTCTGTATCATCTCTGGTGCATGGCAACCATGCCCTTTGACGG 767  
Db 1353 TGTGTGGAGTCTGGGGTCAATTTTATACACACTAGTCAGTGGGCTCACTTCCCTTTGATGG 1412  
QY 768 GCAGGATCATAAACACTGTGTGAAGCAATCAGTAACGGGGCTTACGTAAGCGGCCCAA 827  
Db 1413 GCAAAACCTAAAGAACTGAGAGAGAGAGTATTAAGAGGGGAAATACAGAAATCCCTTCTA 1472

QY 828 GCGCTCGATCGCTGTG---GCCTGATCCGGTGGCTGTTAATGGTGAACCCACCCGTCG 884  
Db 1473 CATGCTTACAGACTGTGAAACCTTCTCAAAACGTTTCTCGTGTCTAAATCAATTAACG 1532  
QY 885 GGCCACACTGGAGGATGTAGCCAGTCAATTTGGTGGTCAA 923  
Db 1533 CGGCACCTTAGACCAATCATGAAGCAGGTGGATCAA 1571

## RESULT 15

US-10-995-561-360  
; Sequence 360, Application US/10995561  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; TITLE OF INVENTION: DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001559  
; CURRENT APPLICATION NUMBER: US/10/995.561  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 360  
; LENGTH: 3514  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-995-561-360

Query Match 9.5%; Score 179.4; DB 7; Length 3514;  
Best Local Similarity 53.1%; Pred. No. 1.5e-29;  
Matches 403; Conservative 1; Mismatches 352; Indels 3; Gaps 1;  
QY 168 CTACGAGTTCCTGGAGCGCTGGCAAGGCGACCTACGGGAAGGTGAAGAGGCACGAGA 227  
Db 813 CTACAGACTGTGAAACAAATCGCAAGGGGAATTTTGCAAAAGTAAATTTGCAAGACA 872  
QY 228 GAGCTCGGGCGTCTGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGCA 287  
Db 873 TATCCTTACAGGAGAGAGTTCGCAATAAATAATTTGCAAAACTCAGTTGAATCAAC 932  
QY 288 GGATCTGCTGCACATACGGAGGAGATTGAGATCATGCTTCACTCAACCCACCCACAT 347  
Db 933 AAGTCTACAAAGCTCTTCAGAGAGTAAAGATAAGAAATTTAAATCATCCCAATAT 992  
QY 348 CATTGCCATCCATGAAGTGTGAGATAGCAGCAAGATTGTGATGTCTATGAGTATGC 407  
Db 993 AGTGAAGTTATTCGARGTCAATTGAAACTGAAAAACAACCTCTACTTAATCATGAATATGC 1052  
QY 408 CAGCCGAGGCGATCTGTATGATTACATCAGTGAAGCGCCACGGCTGAGTGGGAGCG 467  
Db 1053 AAGTGAAGTGAAGTATTTGACTATTTGGTTGCACATGGCAGGATGAAGAAAAGAGC 1112  
QY 468 CAGGCAATTTCTCCGACAGATCGTCTGCCCTGCACTACTGCCACAGAACCGGATCGT 527  
Db 1113 AAGATCTAAATTTAGACAGATTGTCTGCAGTTCAATATCTGCATCAGAAACGGATCGT 1172  
QY 528 TCACCGAGATCTCAAGCTGGAAACATCTTCTAGATGCCAATGGAAACATCAAGATTGC 587  
Db 1173 ACATCGAGACTCAAGGCTGAAATCTATTGTTAGTGGCGATATGAACATTAATAATGC 1232  
QY 588 TGACTTTGGCCTCTCCAACTGTACCAAGGCAAGTTCTCCAGACGTTCTGTGGGAG 647  
Db 1233 AGATTTGGGTTTATGCAATGAAATTTACTGTTGGCGGTAAACTCGACACGTTTGTGGCAG 1292  
QY 648 CCCTCTACGCTCGCTCAGATAGTCAACCGGAAGCCCTATGTGGGCCACAGAGTGA 707  
Db 1293 TCCTCCATACGAGCACTGAGCTCTTCAGGGCAAGAAATATGACGGGCCAGAGTGA 1352  
QY 708 CAGCTGGTCTCTGGGCGTCTCTGTATCATCTCTGGTGCATGGCAACCATGCCCTTTGACGG 767  
Db 1353 TGTGTGGAGTCTGGGGTCAATTTTATACACACTAGTCAGTGGGCTCACTTCCCTTTGATGG 1412  
QY 768 CAGCTGGTCTCTGGGCGTCTCTGTATCATCTCTGGTGCATGGCAACCATGCCCTTTGACGG 767  
Db 1353 TGTGTGGAGTCTGGGGGTCAATTTTATACACACTAGTCAGTGGGCTCACTTCCCTTTGATGG 1412

Qy	768	GCAGGATCATAAACACTGGTGAAGCAATCAGTAAGGGGCTTACCGTGAGCGCCCAA	827
Db	1413	GCATAACCTAAAGGAACCTGAGAGAGAGAGTATTAGAGGGGAAATACAGAAATCCCTTCTA	1472
Qy	828	GCCGTCGATGCCCTGTG---GCCTGATCCGGTGGCTTTAATGGTGAACCCACCCGTCG	884
Db	1473	CATGCTACAGACTGTGAAACCTTCTCAACGTTTCTCGTGTCTAAATCCAATTAAACG	1532
Qy	885	GGCCACACTGGAGGATGTAGCCAGTCATTGGTGGTCAA	923
Db	1533	CGGCACCTCTAGAGCAAAATCATGAAGGACAGGTGATCAA	1571

Search completed: January 25, 2005, 07:13:05  
 Job time : 357.728 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 14:56:28 ; Search time 235.428 Seconds  
(without alignments)  
8761.524 Million cell updates/sec

Title: US-09-980-464-4

Perfect score: 2902

Sequence: 1 cactagtgcattccaagaat.....aaaaaaaaaaaaaaaaaaaaa 2902

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfileseq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2902	100.0	2902	4	US-09-579-664B-4
2	2902	100.0	2902	4	US-10-355-975A-4
3	1372.2	47.3	2501	4	US-09-799-451-215
4	202.6	7.0	2175	4	US-09-984-890-1
5	202.6	7.0	2175	4	US-10-274-194-1
C 6	189.2	6.5	484	4	US-09-270-767-1257
C 7	189.2	6.5	484	4	US-09-270-767-16539
8	186.4	6.4	2908	4	US-09-930-181-1
9	182.4	6.3	2652	3	US-08-557-006C-39
10	182.4	6.3	2761	3	US-08-557-006C-24
11	181.2	6.2	1742	3	US-08-557-006C-38
12	179.8	6.2	2698	2	US-08-677-298-1
13	176.4	6.1	2899	4	US-09-774-528-112
14	174	6.0	3364	4	US-09-930-181-3
15	173	6.0	2334	4	US-10-003-690-3
16	173	6.0	2337	4	US-10-116-326-1
17	173	6.0	2980	4	US-10-003-690-1
18	172.2	5.9	1539	4	US-09-633-328B-1
19	172	5.9	3609	4	US-09-799-875-6
20	172	5.9	5983	4	US-09-799-875-4
21	171.4	5.9	1647	3	US-09-101-146-44
22	165.6	5.7	1929	3	US-09-359-161-4
23	163.6	5.6	1747	3	US-08-557-006C-44
24	160.4	5.5	1507	4	US-09-523-849-4
25	160.4	5.5	1544	4	US-09-734-673-1
26	160.4	5.5	2060	4	US-09-523-849-1
27	158	5.4	1736	3	US-08-557-006C-37
					Sequence 4, Appli
					Sequence 4, Appli
					Sequence 215, App
					Sequence 1, Appli
					Sequence 1, Appli
					Sequence 1257, Ap
					Sequence 16539, A
					Sequence 1, Appli
					Sequence 39, Appl
					Sequence 24, Appl
					Sequence 38, Appl
					Sequence 1, Appli
					Sequence 112, App
					Sequence 3, Appli
					Sequence 1, Appli
					Sequence 1, Appli
					Sequence 6, Appli
					Sequence 4, Appli
					Sequence 44, Appl
					Sequence 4, Appli
					Sequence 44, Appl
					Sequence 4, Appli
					Sequence 1, Appli
					Sequence 1, Appli
					Sequence 37, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-579-664B-4  
; Sequence 4, Application US/09579664B  
; Patent No. 6514719  
; GENERAL INFORMATION:  
; APPLICANT: Immunex Corporation  
; APPLICANT: Bird, Timothy A.  
; APPLICANT: Virca, G. Duke  
; APPLICANT: Martin, Unja  
; APPLICANT: Anderson, Dirk M.  
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES  
; FILE REFERENCE: 2923-A  
; CURRENT APPLICATION NUMBER: US/09/579, 664B  
; CURRENT FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 2902  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-579-664B-4

Query Match	100.0%;	Score 2902;	DB 4;	Length 2902;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2902;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	CAC TAGTGGATCCAAAGAAATTCGGCACGAGGCGTGTCTCGGGTGGCGGTGTGACCTCTGAG	60	
Db	1	CAC TAGTGGATCCAAAGAAATTCGGCACGAGGCGTGTCTCGGGTGGCGGTGTGACCTCTGAG	60	
Qy	61	CCCGCGGCTCAGCGCGCTGTCTACTGTGCGCGACCCACTCCACTCCCGGCTCCCGGCA	120	
Db	61	CCCGCGGCTCAGCGCGCTGTCTACTGTGCGCGACCCACTCCACTCCCGGCTCCCGGCA	120	
Qy	121	CCATGAGTGGTGGGCTTACTCCAGCGCGCGAGCCAGGCTCCCTCGGCTCCGCGCTGG	180	
Db	121	CCATGAGTGGTGGGCTTACTCCAGCGCGCGAGCCAGGCTCCCTCGGCTCCGCGCTGG	180	
Qy	181	CCTCGGAGAGCGCGCGCGCTGTGGCGGACGGGCTCATCAAGTCGCCTAAACCTCTGATGA	240	
Db	181	CCTCGGAGAGCGCGCGCGCTGTGGCGGACGGGCTCATCAAGTCGCCTAAACCTCTGATGA	240	
Qy	241	AGAAGCAGCGGTGAAGCGGCAACCATCAAAACAAACCTGCGGCAACCGCTACGAGTTCC	300	
Db	241	AGAAGCAGCGGTGAAGCGGCAACCATCAAAACAAACCTGCGGCAACCGCTACGAGTTCC	300	
Qy	301	TGGAGACGCTGGGCAAGGCGACCTACGGNAGGTGAAGAGGACGAGAGCTCGGGG	360	
Db	301	TGGAGACGCTGGGCAAGGCGACCTACGGNAGGTGAAGAGGACGAGAGCTCGGGG	360	





Db 1381 CTACCTCGTCAGGGAGGTA CAGGAGGACCTCAGGAACCTCAGACCGTGCTCATCTC 1440  
Qy 1441 CAGGGACGCTGTCCTGCTGTATCCCTGCTCCCAAGAAAGGATCTTTAAGAGTCTC 1500  
Db 1441 CAGGGACGCTGTCCTGCTGTATCCCTGCTCCCAAGAAAGGATCTTTAAGAGTCTC 1500  
Qy 1501 GACAGCGTGAATCTGGTTACTACTCTCTCCAGAGCCAGGAGCTCTGGGGAACCTCTAG 1560  
Db 1501 GACAGCGTGAATCTGGTTACTACTCTCTCCAGAGCCAGGAGCTCTGGGGAACCTCTAG 1560  
Qy 1561 ACGCCAGTGTATGTTTGTAGTGGGACCCCGTGGAGCAAGTCTCCACAGGCTTCAG 1620  
Db 1561 ACGCCAGTGTATGTTTGTAGTGGGACCCCGTGGAGCAAGTCTCCACAGGCTTCAG 1620  
Qy 1621 GGCTCCTCTCCACCGCAAGGGCAATCTCAAACTCAATGGCAAGTCTCCCGCACGCT 1680  
Db 1621 GGCTCCTCTCCACCGCAAGGGCAATCTCAAACTCAATGGCAAGTCTCCCGCACGCT 1680  
Qy 1681 TAGAAGGCATACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCTCTCCCATCTG 1740  
Db 1681 TAGAAGGCATACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCTCTCCCATCTG 1740  
Qy 1741 CAGCCGGCCAGCCGCTCAGGGCTGTGAGTGAGGACAGCATCTGTCTCCGAGT 1800  
Db 1741 CAGCCGGCCAGCCGCTCAGGGCTGTGAGTGAGGACAGCATCTGTCTCCGAGT 1800  
Qy 1801 CCTTTGACCAATGGACTTGGCTGAACTCTTCCGAAACCCCACTGAGGGCTGTGTGT 1860  
Db 1801 CCTTTGACCAATGGACTTGGCTGAACTCTTCCGAAACCCCACTGAGGGCTGTGTGT 1860  
Qy 1861 CTGTGACCACTGAGGGGCTTGCAGCAGCTCTCCTCAGAAAGTCTCAAGCGATGGTGC 1920  
Db 1861 CTGTGACCACTGAGGGGCTTGCAGCAGCTCTCCTCAGAAAGTCTCAAGCGATGGTGC 1920  
Qy 1921 AGGAATCCTTTGGGGATAGCTGCTTTTCTCTGACAGACTGCCAAGAGTGAAGTGCACCT 1980  
Db 1921 AGGAATCCTTTGGGGATAGCTGCTTTTCTCTGACAGACTGCCAAGAGTGAAGTGCACCT 1980  
Qy 1981 ACAGACAAGCCCTAGGAATCTGCTCAAGCTCAGCTAGGAGGAGATGGTCCCTAGT 2040  
Db 1981 ACAGACAAGCCCTAGGAATCTGCTCAAGCTCAGCTAGGAGGAGATGGTCCCTAGT 2040  
Qy 2041 ATGGGTAGGCTCTGAGAGGCTTTCAGAGAAACCTTGGCTGGATTCCTCCAGTGAATA 2100  
Db 2041 ATGGGTAGGCTCTGAGAGGCTTTCAGAGAAACCTTGGCTGGATTCCTCCAGTGAATA 2100  
Qy 2101 GAGTACATCAAGGCTCTAGCTCTGCAAGCCTGACTGAACCTGAAAGATGAGAGAAATCGC 2160  
Db 2101 GAGTACATCAAGGCTCTAGCTCTGCAAGCCTGACTGAACCTGAAAGATGAGAGAAATCGC 2160  
Qy 2161 ATTGATGTGAAAGGAATGGAAACCTTGTGCTCCGAGTCTTATAGTGGGTGGCTGAA 2220  
Db 2161 ATTGATGTGAAAGGAATGGAAACCTTGTGCTCCGAGTCTTATAGTGGGTGGCTGAA 2220  
Qy 2221 GGTGCTTACCTCTTGTGCTAGTGTGACCATGACATTTCCCAACCTGTTCTCTGG 2280  
Db 2221 GGTGCTTACCTCTTGTGCTAGTGTGACCATGACATTTCCCAACCTGTTCTCTGG 2280  
Qy 2281 CTGCACTTCAATAAGTTTCTGTTTCATCAACCCAGGGTTAGAACCTGACTTCCT 2340  
Db 2281 CTGCACTTCAATAAGTTTCTGTTTCATCAACCCAGGGTTAGAACCTGACTTCCT 2340  
Qy 2341 GGGAGGTAATGTAGTGTGCTGCTGATTTAGAGGAAACAGCTCTGGTTTCCATCT 2400  
Db 2341 GGGAGGTAATGTAGTGTGCTGCTGATTTAGAGGAAACAGCTCTGGTTTCCATCT 2400  
Qy 2401 CTGCTGTGTGATCTCAAAAGACTGGGAAAGCTCGGACCGCTGTTGACTTCATCTCA 2460  
Db 2401 CTGCTGTGTGATCTCAAAAGACTGGGAAAGCTCGGACCGCTGTTGACTTCATCTCA 2460  
Qy 2461 GGGGACCAAGTGGCCCTGGACCCCATCTTATGATCTCAGACCTTGAACCTTGAAGCTGT 2520

Db 2461 GGGGACCAAGTGGCCCTGGACCCCATCTTATGATCTCAGACCTTGAACCTTGAAGCTGT 2520  
Qy 2521 CCTAGTACCAAGATGTGGATGGATGCTCTGTTTCTCAGGCCAACGGGACCTAGAAATGTC 2580  
Db 2521 CCTAGTACCAAGATGTGGATGGATGCTCTGTTTCTCAGGCCAACGGGACCTAGAAATGTC 2580  
Qy 2581 TGACTTATTTATTTTGTGATTTCTCAGCTCTGTTTCTGTTTGTGTTTGTGTTTGTGTTT 2640  
Db 2581 TGACTTATTTATTTTGTGATTTCTCAGCTCTGTTTCTGTTTGTGTTTGTGTTTGTGTTT 2640  
Qy 2641 TTTTGTGTTTAAAGTGAATTTTCTGCTTTTCAATTAATGTGAATGATGTCCTCTATTC 2700  
Db 2641 TTTTGTGTTTAAAGTGAATTTTCTGCTTTTCAATTAATGTGAATGATGTCCTCTATTC 2700  
Qy 2701 TCACCTGTGCCATGAAGTTTATGTACAGAAAGTATTTGGCAATGATGTCCTCTATTC 2760  
Db 2701 TCACCTGTGCCATGAAGTTTATGTACAGAAAGTATTTGGCAATGATGTCCTCTATTC 2760  
Qy 2761 AAGGGGGTGGGGGGCTTTTCAAAATGATGTTGAGCACTGCTGGATTTGATGTTGTTT 2820  
Db 2761 AAGGGGGTGGGGGGCTTTTCAAAATGATGTTGAGCACTGCTGGATTTGATGTTGTTT 2820  
Qy 2821 GTCCCTTCACACCAAGGCTGGCCACCTCCCTCATCTTCATCTGTGGCCAAAAA 2880  
Db 2821 GTCCCTTCACACCAAGGCTGGCCACCTCCCTCATCTTCATCTGTGGCCAAAAA 2880  
Qy 2881 AAAAAAAAAAAAAAAAAAAAAA 2902  
Db 2881 AAAAAAAAAAAAAAAAAAAAAA 2902

## RESULT 3

US-09-799-451-215  
; Sequence 215, Application US/09799451  
; Patent No. 6783969  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Zhang, Jie  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Ma, Yundong  
; APPLICANT: Yamazaki, Victoria  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Ghosh, Reena  
; APPLICANT: Dmanac, Radoje T.  
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and  
; FILE REFERENCE: 803  
; CURRENT APPLICATION NUMBER: US/09/799,451  
; CURRENT FILING DATE: 2001-03-05  
; NUMBER OF SEQ ID NOS: 948  
; SOFTWARE: pt FL\_genes Version 2.0  
; SEQ ID NO 215  
; LENGTH: 2501  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (106)...(1989)  
US-09-799-451-215

Query Match 47.3%; Score 1372.2; DB 4; Length 2501;  
Best Local Similarity .81.6%; Pred. No. 0;  
Matches 1649; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

QY 82 CTACTGCTGCGCCGACCACTCCACTCGCGGTCCCCGCAACATGAGTGGTGGCTTAC 141  
DB 65 CTATTGATTCCCTGCGCCCTTGTCTCACCTCTGCTCGCATGAGTGGTGGTTCG 124  
QY 142 TCAGCGCCGACGACGCTCCCTGCGGCTCGGCTCGGCTCGGAGAGCGCCGCGC 201  
DB 125 CGCGGCGCTCGGCGCCCACTCCCTCGGCGCAGAGCTA-----GCCGCGCGC 172  
QY 202 TGGCGGACGGGCTCATCAAGTCGGCTAAACCTCTGATGAAGACAGCGCGTGAAGCGC 261  
DB 173 TGGCGGAGGGCTGATCAAGTCGCGCCCAAGCCCTAATGAAGACAGCGCGTGAAGCGC 232  
QY 262 ACCATCAAAACAACTCGGCAACCGCTAAGAGTTCTTGAGACGCTGGCAAGGGCA 321  
DB 233 ACCACCAAGACCAACTCGGCAACCGCTAAGAGTTCTTGAGACGCTGGCAAGGGCA 292  
QY 322 CCTACGGGAAGGTGAAGAGGACGACGAGAGCTCGGCGGCTGCTGGCCATCAAGTCCA 381  
DB 293 CCTACGGGAAGGTGAAGAGGCGCGGAGAGCTCGGCGGCTGCTGGCCATCAAGTCAA 352  
QY 382 TCAGGAAGCAAAATCAAGATGAGCAGGATCTGTCACATACGAGGAGGATTCAGA 441  
DB 353 TCGGAAGGCAAAATCAAGATGAGCAGGATCTGATGCATATACGAGGAGGATTCAGA 412  
QY 442 TCATGCTTCACTCAACCCCACTCAATTCGCTCCATCCATGAGTGTGTTGAGAAATAGCA 501  
DB 413 TCATGCTCACTCAACCCCACTCAATTCGCTCCATCCATGAGTGTGTTGAGAAATAGCA 472  
QY 502 GCAAGATTGATTGTCATGAGGATGCGAGCGGAGGATCTGATGATTAATCAATCAGTG 561  
DB 473 GCAAGATCGTGATCGTATGAGGATGCGAGCGGAGGATCTGATGATTAATCAATCAGTG 532  
QY 562 AGCGGCCACGGCTGAGTGGGAGCGCAGGCAATTCCTCGACAGATCGTGTCTGCC 621  
DB 533 AGCGGCCACGGCTGAGTGGGAGCGCAGGCAATTCCTCGCGCAGATCGTGTCTGCC 592  
QY 622 TGCACTACTGCCACGAGCGGATCGTTCCACGAGATCTCAAGCTGGAAATCATCTTC 681  
DB 593 TGCACTATTGCCATCAGACAGAGTGTTCACCGAGATCTCAAGCTGGAGAAATCTCTCT 652  
QY 682 TAGATGCAATGGAACATCAAGATTGCTGACTTTGGCTCTCAACCTGTACACAAAG 741  
DB 653 TGGATGCAATGGAATATCAAGATTGCTGACTTGGGTCTCTCAACCTCTACCATCAAG 712  
QY 742 GCAAGTTCCTCGACGCTTCTGTTGGAGCCCTCTACGCGCTCGCTGAGATAGTCAAG 801  
DB 713 GCAAGTTCCTCGACGATCTGTTGGAGCCCTCTATGCTCGCCAGAGATTGTCAATG 772  
QY 802 GGAAGCCCTATGTGGGCCAGAGGTGGACAGTGGTCTCTGGGCGTTCTCTGTACATCC 861  
DB 773 GGAAGCCCTACAGGCCAGAGGTGGACAGTGGTCTCTGGGCTCTCTCTACATCC 832  
QY 862 TGTGATGTCACATGCCCTTTGACGGGAGGATCATAAACAATGTTGAAGCAATCA 921  
DB 833 TGGTGCATGTCACATGCCCTTTGATGGGATGACCAATAGATCTTAGTGAACACAGATCA 892  
QY 922 GTAAACGGGCTTACCGTGAAGCGCCCAAGCGCTCGATGCTGTGGCTGATCGGTGGC 981  
DB 893 GCAACGGGCTTACCGGAGCCCACTAAACCTCTGATGCTGTGGCTGATCGGTGGC 952  
QY 982 TGTTAATGTGTAACCCCACTCGGCGCACACTGGAGGATGAGCCAGTCAATTTGGTGGG 1041  
DB 953 TGTGATGTGTAACCCCACTCGGCGCACCTTGGAGGATGAGCCAGTCACTTGGTGGG 1012  
QY 1042 TCAACTGGGTTTACACCAACCGGAGTCGGGAAACAGGAGCCCTCGGTAGGGTGGGCACC 1101  
DB 1013 TCAACTGGGCTACGCCCAACCGAGTGGAGAGCAGGAGGCTCGCATGAGGGTGGGCACC 1072  
QY 1102 CTAGTGTGATCTTTGGCCGGCTCCATGCGGAGTGGTTACGTGCTCTCGGCGCC 1161  
DB 1073 CTGGAGTGAATCTGCGCGCGCTCCATGCTGACTGGCTCGCGCGTTCCTCCGCGCC 1132

QY 1162 TCCTGGAGAAATGAGGCAAGTGTGACAGTCTTCTCAAGCAGACAGTCCGCGAGGTGAA 1221  
DB 1133 TCCTGGAGAAATGAGGCAAGTGTGACAGTCTTCTCAAGCAGACAGTCCGCTGGTGGGAA 1192  
QY 1222 GCACTGTACTGGGCTGGAGGCGCAATTTCTTTAAGAGTCCGAAAGAGAGATGACA 1281  
DB 1193 GCACCACTCCCTGGCTGGAGCGCGCATTTGCTCAAGAAAGTCCCGAAGAGAGATGACA 1252  
QY 1282 TGGCTCAAAATCTGCAAGGTGACCGGCTGAGGATACCTTCTCTCGCCCTGGCAAGAGCA 1341  
DB 1253 TGGCCAGTCTCTCCACAGTGAACGGCTGATGACATCGCCATCGCCCTGGCAAGAGCA 1312  
QY 1342 GCCTTAAGCTTCCGAAAGGCAATTTCTCAAGAAAGTCTCTACCTCTGTCAGGGAGGTAC 1401  
DB 1313 ACCTAAGCTGCCAAAGGGCATTTCTCAAGAAAGGTGTGAGCTCTGCAAGAGGGGTAC 1372  
QY 1402 AGAGGACCTTCAGGAACCTCAGACCGGTGCTGATATCTCAGGGCAGCTGTCCCTGCTG 1461  
DB 1373 AGAGGACCTTCGCGAGCTCAGCCCAATCCCTCGAGCCCGAGGCGAGGCTGCC----- 1426  
QY 1462 TATCCCTGCTCCCAAGGAAAGGCATCTTAAAGAGTCTCGACAGCGTGAATCTGTTTACT 1521  
DB 1427 ---CCCTGCTCCCAAGAGGGCATTTCTCAAGAAAGCCCGACAGCGGAGTCTGGCTACT 1483  
QY 1522 ACTCTCTCCAGACCCGAGCGAGTCTGGGGAATCTTTAGACGCGAGTGTGTTTGTGA 1581  
DB 1484 ACTCTCTCCGAGCCCGAGTGAATCTGGGAGCTCTTGGACGAGCGGAGCTGTTTGTGA 1543  
QY 1582 GTGGGACCCCGTGGAGCAGAGTCTCCAAGGTTTCAGGGCTCTCTCTCACCGCAAGG 1641  
DB 1544 GTGGGATCCCAAGGAGCAGAGCTCCGCAAGCTTCAGGGCTGCTCTCTCATCGCAAG 1603  
QY 1642 GCATCTCAAACTCAATGGCAAGTCTCCCGCAGCAGCTTTAGAGGACATACCCCTAGCA 1701  
DB 1604 GCATCTCAAACTCAATGGCAAGTCTCCCGCAGCAGCTTTGGAGCTCGCGCCCGCCACCA 1663  
QY 1702 CCTTGGCTCCCTGGACCAACTGGCTCTCCCATCTGCGAGCCCGCGCCGAGCGCCCT 1761  
DB 1664 CCTTGGCTCCCTGGATGAATCTGCGCCACCTCGCCCTCGCCCGGCGCAGCGACCT 1723  
QY 1762 CAGGGCTGTGAGTGGAGCAGCATCTGTCTCTCGAGTCTTTGACAAATTTGACTTGC 1821  
DB 1724 CAGGGCTGTGAGCAGGACAGCATCTGTCTCTGAGTCTTTGACAGCTGACTTGC 1783  
QY 1822 CTGAACTCTTCCGAAACCCCACTGAGGGCTGTGTCTGTGGAGCAACTGAGGGGC 1881  
DB 1784 CTGAACTCTTCCGAGACCCCACTGCGGGCTGTGTCTGTGGAGCAACTCAGGGGC 1843  
QY 1882 TTGAGCAGCTCCCTCAGAG-----GTCTGAAGCATGGTGGCAGGAATCCT 1929  
DB 1844 TTGAGGAGCCCTCAGAGGGCCCTGGAAGCTGCCTGAGCGCTGGCGCAGGATCCTT 1903  
QY 1930 TGGGGATAGTCTCTTCTGACAGCTGCCAAGAGTGTGACTGAGCTTACAGACAAG 1989  
DB 1904 TGGGGACAGCTCTTCTTCTGACAGCTGCCAGGAGTGTGACAGCACTTACGACAGG 1963  
QY 1990 CCTAGATCTCTCAAGCTCAGCTGAGGAGGAGATGGTGGCTTGTATGGGCTAG 2049  
DB 1964 CACTGAGGGTCTGCTCAAGAGCTCACCTGAGTGGAGTAGGCAATTCGCCAG-CCCGGTGAG 2022  
QY 2050 GCTCTGAGAGGGTTTGGAGAGGAACTCCCTGGGTCGATTCCT 2090  
DB 2023 GCTCTGAGTGCAGCTGGTTGGACCCCGGAGGAGATGCTT 2063

## RESULT 4

US-09-984-890-1

; Sequence 1, Application US/09984890

; Patent No. 6492156

; GENERAL INFORMATION:

; APPLICANT: YAN, Chunhua et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES



; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001306  
; CURRENT APPLICATION NUMBER: US/09/984,890  
; CURRENT FILING DATE: 2001-10-31  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2175  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-984-890-1

Query Match 7.0%; Score 202.6; DB 4; Length 2175;  
Best Local Similarity 54.9%; Pred. No. 4e-42;  
Matches 422; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

```
QY 290 CTACGAGTTCCTGGAGACGCTGGGCAAGGCACTACGGAAGGTGAAGAGGCACGAGA 349
    |||||
Db 156 CTACCGGCTCCTCAAGACCATGGCAAGGTAAATTTTGGCAAGGTGAAGTTGGCCGACA 215

QY 350 GAGCTCGGGCGTCTGGTGCCATCAAGTCCATCAGGAAAGACAAATCAAGATGAGCA 409
    |||||
Db 216 CATCTGACTGGGAAAGAGTAGCTGTGAAGATCATTCAGAGACTCAACTGAATCCTC 275

QY 410 GGATCTGTCACATCGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCACAT 469
    |||||
Db 276 CAGCTCCAGAAACTATTCGCGAAGTAAGAAATTAAGGTTTGAATCATCCAAAT 335

QY 470 CATTTGCCATCCATGAAGTGTTCAGAAATAGCAGCAAGATTGTGATTCATGAGTATGC 529
    |||||
Db 336 AGTTAAATATTATTTGAAGTGAATTCGCGAAGTAAAGATTAATGAAGTTTGAATCATCCAAAT 335

QY 530 CAGCCGAGGCGATCTGTATGATTACATCAGTGAGCGGCCACGCTCAGTGAGCGGACGC 589
    |||||
Db 396 TAGTGGCGAGAGGTATTGATTACCTAGTGGCTCATGGCAGGATGAAGAAAAGGCG 455

QY 590 CAGGCATTTCTCCGACAGATCTGTCTGCGCTGCATCTGCGACACTGCGCACCAGAACGGGATCGT 649
    |||||
Db 456 TCAGGCCAATTCGCGCAGGTAGTGTCTGTGTGCGACTGTGTCAACGAAAGTTATTGT 515

QY 650 TCACCGAGATCTCAAGCTGGAACAATCTCTTAGATGCAATGGAACAATCAAGATTGC 709
    |||||
Db 516 CCATAGAGACTTAAAGGCGAAGAACTGCTCTTGGATGCTGATGAACATCAAGATTGC 575

QY 710 TGACTTTGGCTTCAAGCTGGAACAATCTCTTAGATGCAATGGAACAATCAAGATTGC 769
    |||||
Db 576 AGACTTTGGCTTCAAGCTGGAACAATCTCTTAGATGCAATGGAACAATCAAGATTGC 635

QY 770 CCCTCTTACGCTCGCTCAGATAGTCAACGGGAGCCCTATGTGGGCCACAGAGTGA 829
    |||||
Db 636 TCCCTCTTATGCTGCCCCAGACTCTTCCAGGGCAAAATATGATGGACCGAGGTGA 695

QY 830 CAGCTGGTCTCTGGGCGTCTCTCTGTACATCTCTGTGATGCAATGGAACAATCAAGATTGC 889
    |||||
Db 696 TGTGTGGAGCTTAGGAGTTATCTCTATACACTGCTCAGCGGATCCCTGCTTTGATGG 755

QY 890 GCAGGATCAAAACACTGTGTGAAGCAATCAGTAACGGGGCTTACCGTGAAGCCGCCAA 949
    |||||
Db 756 ACAGAACCTCAGGAGCTGGGAAACGGGTACTGAGGGGAAATATACCGTATCCATTCTA 815

QY 950 GCCGTCCGATGCTGTG---GCCTGATCCGGTGGCTTAAATGGTGAACCCCGCTCG 1006
    |||||
Db 816 CATGTCCAGGACTGTGAAACCTGTCTTAAGAAATTTCTCATTTCTTAATCCGAGAGAG 875

QY 1007 GGCCACACTGGAGGATGTAGCCAGTCAATTTGGTGGGTCAACTGGGGTTAC 1055
    |||||
Db 876 AGGCACTTTAGAGCAAAATCATGAAGATCGATGATGAATGTGGGTAC 924
```

RESULT 5

US-10-274-194-1

; Sequence 1, Application US/10274194

; Patent No. 6705511

; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CL001306DIV  
; CURRENT APPLICATION NUMBER: US/10/274,194  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2175  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-274-194-1

Query Match 7.0%; Score 202.6; DB 4; Length 2175;  
Best Local Similarity 54.9%; Pred. No. 4e-42;  
Matches 422; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

```
QY 290 CTACGAGTTCCTGGAGACGCTGGGCAAGGCACTACGGAAGGTGAAGAGGCACGAGA 349
    |||||
Db 156 CTACCGGCTCCTCAAGACCATGGCAAGGTAAATTTTGGCAAGGTGAAGTTGGCCGACA 215

QY 350 GAGCTCGGGCGTCTGGTGCCATCAAGTCCATCAGGAAAGACAAATCAAGATGAGCA 409
    |||||
Db 216 CATCTGACTGGGAAAGAGTAGCTGTGAAGATCATTCAGAGACTCAACTGAATCCTC 275

QY 410 GGATCTGTCACATCGGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCACAT 469
    |||||
Db 276 CAGCTCCAGAAACTATTCGCGAAGTAAAGATTAATGAAGTTTGAATCATCCAAAT 335

QY 470 CATTTGCCATCCATGAAGTGTTCAGAAATAGCAGCAAGATTGTGATTCATGAGTATGC 529
    |||||
Db 336 AGTTAAATATTATTTGAAGTGAATTCGCGAAGTAAAGATTAATGAAGTTTGAATCATCCAAAT 335

QY 530 CAGCCGAGGCGATCTGTATGATTACATCAGTGAGCGGCCACGCTCAGTGAGCGGACGC 589
    |||||
Db 396 TAGTGGCGAGAGGTATTGATTACCTAGTGGCTCATGGCAGGATGAAGAAAAGGCG 455

QY 590 CAGGCATTTCTCCGACAGATCTGTCTGCGCTGCATCTGCGACACTGCGCACCAGAACGGGATCGT 649
    |||||
Db 456 TCAGGCCAATTCGCGCAGGTAGTGTCTGTGTGCGACTGTGTCAACGAAAGTTATTGT 515

QY 650 TCACCGAGATCTCAAGCTGGAACAATCTCTTAGATGCAATGGAACAATCAAGATTGC 709
    |||||
Db 516 CCATAGAGACTTAAAGGCGAAGAACTGCTCTTGGATGCTGATGAACATCAAGATTGC 575

QY 710 TGACTTTGGCTTCAAGCTGGAACAATCTCTTAGATGCAATGGAACAATCAAGATTGC 769
    |||||
Db 576 AGACTTTGGCTTCAAGCTGGAACAATCTCTTAGATGCAATGGAACAATCAAGATTGC 635

QY 770 CCCTCTTACGCTCGCTCAGATAGTCAACGGGAGCCCTATGTGGGCCACAGAGTGA 829
    |||||
Db 636 TCCCTCTTATGCTGCCCCAGACTCTTCCAGGGCAAAATATGATGGACCGAGGTGA 695

QY 830 CAGCTGGTCTCTGGGCGTCTCTCTGTACATCTCTGTGATGCAATGGAACAATCAAGATTGC 889
    |||||
Db 696 TGTGTGGAGCTTAGGAGTTATCTCTATACACTGCTCAGCGGATCCCTGCTTTGATGG 755

QY 890 GCAGGATCAAAACACTGTGTGAAGCAATCAGTAACGGGGCTTACCGTGAAGCCGCCAA 949
    |||||
Db 756 ACAGAACCTCAGGAGCTGGGAAACGGGTACTGAGGGGAAATATACCGTATCCATTCTA 815

QY 950 GCCGTCCGATGCTGTG---GCCTGATCCGGTGGCTTAAATGGTGAACCCCGCTCG 1006
    |||||
Db 816 CATGTCCAGGACTGTGAAACCTGTCTTAAGAAATTTCTCATTTCTTAATCCGAGAGAG 875

QY 1007 GGCCACACTGGAGGATGTAGCCAGTCAATTTGGTGGGTCAACTGGGGTTAC 1055
    |||||
Db 876 AGGCACTTTAGAGCAAAATCATGAAGATCGATGATGAATGTGGGTAC 924
```

## RESULT 6

US-09-270-767-1257/c  
; Sequence 1257, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1257  
; LENGTH: 484  
; TYPE: DNA  
; ORGANISM: *Drosophila melanogaster*  
US-09-270-767-1257

Query Match 6.5%; Score 189.2; DB 4; Length 484;  
Best Local Similarity 64.3%; Pred. No. 5.2e-39;  
Matches 284; Conservative 0; Mismatches 158; Indels 0; Gaps 0;  
QY 355 CGGGCGCTCTGGTGGCCATCAAGTCCATCAGGAAAGCAAAATCAAGATGAGCAGGATC 414  
Db 442 CGGGCAGGAGTGGCTATCAAAACCATCAAGAAAGTCAAGATCGAGGCCGCGGATT 383  
QY 415 TGCTGCACATACGAGGAGATTGAGATCATGTCTTCACTCAACCCACCCACATCATTTG 474  
Db 382 TGGTGGCATCCGTCCGGAGTGCAGATTATGAGTCACTGATGATCATCCCAATCATCC 323  
QY 475 CCATCCATGAAGTGTGGAATAGCAGCAAGATTGTGATTGTGCTCATGGATGCTTCCACC 534  
Db 322 ACATCTACGAAGTATTGAGAATCGTGAGAAATGGTGTAGTCAATGGAATTTGCCGCTG 263  
QY 535 GAGCGATCTGTATGATTACATCAGTGAGCGGCCACGGCTGAGTGAGCGGAGCCAGGC 594  
Db 262 GCGCGAGCTCTACGACTATCTGTCTGAAGGAAGGTTCTCACCGAGGAGGCGGAGAC 203  
QY 595 ATTTCTCCGACAGATCGTGTCTGCCCTGCCTACTCTCCACCCAGAACGGGATGTTCCACC 654  
Db 202 GCATCTCCGCCAGGTGGCCCGCTTACTACTGTCAAGCAACAAGATCTGCCATC 143  
QY 655 GAGATCTCAAGCTGGAAAACATCTTCTAGATGCCAATGGAACATCAAGATTGCTGACT 714  
Db 142 GCGATCTCAAGCTGGAGACATCTCTGTCGACGAGAAGGCAATGCTAAGATTGCTGATT 83  
QY 715 TTGGCTCTCCAACTGTATACCAAAAGGCAAGTTCTCTCAGAGCTTCTGTGGAGGCCCTC 774  
Db 82 TTGGGTTGTGAATGTGTTTGTATGACGACGACTGTCTGGGCACTTTTGGGGTTCCCCAC 23  
QY 775 TCTACGCTCGCTGAGATAGT 796  
Db 22 TCTATGCTCGCCGGAATTTGT 1

## RESULT 7

US-09-270-767-16539/c  
; Sequence 16539, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16539  
; LENGTH: 484  
; TYPE: DNA  
; ORGANISM: *Drosophila melanogaster*  
US-09-270-767-16539

## Query Match

6.5%; Score 189.2; DB 4; Length 484;  
Best Local Similarity 64.3%; Pred. No. 5.2e-39;  
Matches 284; Conservative 0; Mismatches 158; Indels 0; Gaps 0;  
QY 355 CGGGCGCTCTGGTGGCCATCAAGTCCATCAGGAAAGCAAAATCAAGATGAGCAGGATC 414  
Db 442 CGGGCAGGAGTGGCTATCAAAACCATCAAGAAAGTCAAGATCGAGGCCGCGGATT 383  
QY 415 TGCTGCACATACGAGGAGATTGAGATCATGTCTTCACTCAACCCACCCACATCATTTG 474  
Db 382 TGGTGGCATCCGTCCGGAGTGCAGATTATGAGCTCAGTGCATCATCCCAATCATCC 323  
QY 475 CCATCCATGAAGTGTGGAATAGCAGCAAGATTGTGATTGTGCTCATGGATGCTTCCACC 534  
Db 322 ACATCTACGAAGTATTGAGAATCGTGAGAAATGGTGTAGTCAATGGAATTTGCCGCTG 263  
QY 535 GAGCGATCTGTATGATTACATCAGTGAGCGGCCACGGCTGAGTGAGCGGAGCCAGGC 594  
Db 262 GCGCGAGCTCTACGACTATCTGTCTGAAGGAAGGTTCTCACCGAGGAGGCGGAGAC 203  
QY 595 ATTTCTCCGACAGATCGTGTCTGCCCTGCCTACTCTCCACCCAGAACGGGATGTTCCACC 654  
Db 202 GCATCTCCGCCAGGTGGCCCGCTTACTACTGTCAAGCAACAAGATCTGCCATC 143  
QY 655 GAGATCTCAAGCTGGAAAACATCTTCTAGATGCCAATGGAACATCAAGATTGCTGACT 714  
Db 142 GCGATCTCAAGCTGGAGACATCTCTGTCGACGAGAAGGCAATGCTAAGATTGCTGATT 83  
QY 715 TTGGCTCTCCAACTGTATACCAAAAGGCAAGTTCTCTCAGAGCTTCTGTGGAGGCCCTC 774  
Db 82 TTGGGTTGTGAATGTGTTTGTATGACGACGACTGTCTGGGCACTTTTGGGGTTCCCCAC 23  
QY 775 TCTACGCTCGCTGAGATAGT 796  
Db 22 TCTATGCTCGCCGGAATTTGT 1

## RESULT 8

US-09-930-181-1  
; Sequence 1, Application US/09930181  
; Patent No. 6455292  
; GENERAL INFORMATION:  
; APPLICANT: Origene Technologies  
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas  
; FILE REFERENCE: 16U 101 V1  
; CURRENT APPLICATION NUMBER: US/09/930,181  
; CURRENT FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 2908  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (106)..(2112)  
US-09-930-181-1

## Query Match

6.4%; Score 186.4; DB 4; Length 2908;  
Best Local Similarity 53.5%; Pred. No. 7.2e-38;  
Matches 413; Conservative 0; Mismatches 356; Indels 3; Gaps 1;  
QY 288 CGCTACGAGTCTCTGGAGACGCTGGGCAAGGGCACCTACGGGAAGGTGAAGAGGCACGA 347  
Db 157 CCTACCGGTGGAGAAGACGCTGGGCAAGGGCACAGGTCGTGGTGAAGCTGGGGGTT 216  
QY 348 GAGAGCTCGGGGCGTCTGGTGGCCATCAAGTCCATCAGGAAAGCAAAAATCAAGATGAG 407  
Db 217 CACTGGTCACTCGCCAGAGGTGGCCATCAAGATCGTCAACCGTGAGAGCTCAGCGAG 276  
QY 408 CAGGATCTGTGCACATACGAGGAGGAGATTGAGATCATGTCTTCACTCAACACCCCCAC 467  
Db 277 TCGGTCTGTGAAGGTGGAGCGGAGATCGCGATCTCTGAAGCTCATTTGAGCAGCCCCAC 336

QY 468 ATCAATTGCCATCATGAAGTGTGAGATAGCAGCAAGTGTGATGTGATGAGTAT 527  
Db 337 GTCTAAAGCTGCACGACGTTTATGAAAAAATAATTTGTACCTGTGCTAGAACAC 396  
QY 528 GCCAGCGGAGGCGATCTGTATGATTAATCATCAGTGAGCGGCCACCGCTGAGTGAGCGGAC 587  
Db 397 GTGTCAAGTGTGAGCTCTTCGACTACCTGGTGAAGAGGGGAGGCTGACCCCTAAGGAG 456  
QY 588 GCAGGCGATTTCTTCGGACAGATCGTGTCTGCCCTGCACTACTGCCACCAAGAAACCGGATC 647  
Db 457 GCTCGGAAGTCTTCGGCAGATCATCTCTGCGCTGGACTTCTGCCACAGCCACTCCATA 516  
QY 648 GTTACCGGAGATCTCAAGCTGGAAGAAACATCTCTCTAGATGCCAATGGAACATCAAGATT 707  
Db 517 TGCCACAGGAGTCTGAACCTGAAACCTCTCTGCGACAGAGAAACAAACATCCGCAATC 576  
QY 708 GCTGACTTTGCGCTCTCCAACTGTACCAAGAGCAAGTTCTCCAGACGTTCTGTGGG 767  
Db 577 GCAGACTTTGGCATGGCGTCCCTGCAGGTTGGCGACAGCCCTGTTGGAGCAGCTGTGGG 636  
QY 768 AGCCCTCTCAAGCCTCGCTGAGATAGTCAACGGGGAAGCCCTATGTGGGCCCGCAGAGTG 827  
Db 637 TCCGCCCACTAGCCTGCCCGAGGTGATCCGGGGGAGAGATATGACGGCCCGGAAGGCG 696  
QY 828 GACAGCTGTCTCTGGCGCTTCTCTGTACATCTCTGTGTCATGGCACCATTGCCCTTTGAC 887  
Db 697 GACGTGTGAGCTGCGCGCTCATCTCTGCTGCCCTTGTGTGGGGGCTCTGCCCTTCGAC 756  
QY 888 GGGCAGGATCATAAACACTGGTGAAGCAAAATCAGTAAACGGGGCTTACCGTGAGCGGCC 947  
Db 757 GATGACAACTTGGACAGCTGCTGGAGAAGGTGAAGCGGGCGTGTCCACATGCCGCAC 816  
QY 948 ---AAGCCGCTCGATGCTGTGGCCTGATCCGCTGTGTTAATGTGGAACCCACCCGT 1004  
Db 817 TTTATCCCGCGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGAGCGCGCACGC 876  
QY 1005 CGGCGCACCTGGAGGATGAGCCAGTCTGTGGGTCACTGGGGTTACA 1056  
Db 877 CGCCTCAGCTAGAGCAATTCAGAAAAACATATGGTATATAGGGGGCAAGA 928

## RESULT 9

US-08-557-006C-39  
; Sequence 39, Application US/08557006C  
; Patent No. 6258547  
; GENERAL INFORMATION:  
; APPLICANT: Beri, Rajindar K.  
; APPLICANT: Carling, David  
; APPLICANT: Forster, Robert A.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
; FILE REFERENCE: NGAP/PHM37588/UST  
; CURRENT APPLICATION NUMBER: US/08/557,006C  
; CURRENT FILING DATE: 1996-03-06  
; PRIOR APPLICATION NUMBER: PCT/GB94/01093  
; PRIOR FILING DATE: 1994-05-20  
; PRIOR APPLICATION NUMBER: GB 9310489.1  
; PRIOR FILING DATE: 1993-05-21  
; PRIOR APPLICATION NUMBER: GB 9318010.7  
; PRIOR FILING DATE: 1993-08-31  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 39  
; LENGTH: 2652  
; TYPE: DNA  
; ORGANISM: Rat  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (1) .. (1747)  
; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -  
; OTHER INFORMATION: fragment begins at nucleotide 24 and ends with  
; OTHER INFORMATION: nucleotide 1765  
US-08-557-006C-39

Query Match 6.3%; Score 182.4; DB 3; Length 2652;  
Best Local Similarity 53.0%; Pred. No. 7.4e-37;  
Matches 437; Conservative 0; Mismatches 381; Indels 6; Gaps 2;  
QY 242 GAAGCAGGCGGTGAAGCGGCACCATCACAAAACAAACCTGCGGCACCGCTACGAGTTCT 301  
Db 4 GAACATGGCTGAGAGCAGAGCAGACGCGGCGTGTGAAGATCGGACACTACGTGCTGGG 63  
QY 302 GGAGACGCTGGGCAAGGGCACTTACGGGAAGGTGAAGAGGACGAGAGAGTCG---GG 358  
Db 64 GGACACCCCTGGGCGTCGGCACCTTCGGCAAAAGTGAAGATTGGAGAACATCAATTCACAG 123  
QY 359 GGGTCTGTGGCCATCAAGTCCATCAGGAACACAAAATCAAAGATGAGCAGGATCTGCT 418  
Db 124 CCATAAAGTGGCAGTTAAGATCTTAATAGACAGAGATTCGAGTTTAGATGTTGTTGG 183  
QY 419 GCACATACGAGGAGATTGAGATCATGTCTTCACTCAACACCCGCCACATCATTTGCCAT 478  
Db 184 AAAAATAAAACAGAGAAATTTCAAAATCTTAAACTCTTCGTCATCTCTCATATATCAAACT 243  
QY 479 CCATGAAGTGTGAGATAGCAGCAAGATTCTGATTGTTCATGGAGTATGCCAGCGGAG 538  
Db 244 CTACCAAGTGATCAGCACTCCAAACAGACTTTTTTATGGTAATGGAAATATGTGTCTGAGG 303  
QY 539 CGATCTGTATGATTACATCAGTGAGCGGCCACGGCTGAGTGAGCGGAGCGCCAGGCAATTT 598  
Db 304 TGAAATTTGCTGACTACTCTGTAAACACGGGAGGGTTGAAGAGGTGGAAGCTCGCCGGCT 363  
QY 599 CTTCCGACAGATCGTGTCTGCCCTGCATCTACTGCCACAGAAACGGGATCGTTCCACGAGA 558  
Db 364 CTTCCAGCAGATCTCTGTCTGCCCTGACTACTGTACAGGCACATGGTTGTCCACAGGGA 423  
QY 659 TCTCAGCTGGAAGAACATCTTCTAGATGCCAATGGAACATCAAGATTCTGACTTTGG 718  
Db 424 CTTGAAGCCAGAGAACGTTGTGTGGACGCCCAGATGAATGCTAAGATAGCTGACTTTCG 483  
QY 719 CTTCTCCAACTGTACCAAAAGGCAAGTTCTCTCCAGACGTTCTGTGGGAGCGCTCTCTA 778  
Db 484 ACTCTCTAATATGATGCTCAGATGGTGAATTTCTACGAATAGCTGTGGATCGCAAAATTA 543  
QY 779 CGCCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCGCAGAGTGAGAGCTGGTC 838  
Db 544 TGACGACCGGAGGTCTATCTCAGGAAGGCTGTATGCGGGTCTCTGAGGTTGATATCTGGAG 603  
QY 839 TCTGGGGTCTCTCTGTACATCTGTGTCATGCGCACCATGCCCTTTGACGGGAGGATCA 898  
Db 604 CTGTGGTGTATCTCTGTATGCCCTTCTCTGTGGCACCCCTCCCTTCGACGATGAGCAGCT 663  
QY 899 TAAACACACTGGTGAAGCAATCAGTAACGGGGCTTACCGTGAGCGCCGCCAAGCGCTCGGA 958  
Db 664 GCCTACGCTCTTTAAGAAAGATCCGAGGGGTGTGTTCTACATCCCGAGTATCTCAACCG 723  
QY 959 TGCTGTGGC---CTGATCCGGTGGCTGTAAATGTGGGAACCCACCCGCTGGGCCACACT 1015  
Db 724 TTCTATTGCCACTCTCTGTATGCACATGCTGCAAGTGGAGCCCTCTGAAGCAGCAACTAT 783  
QY 1016 GGAGGATGAGCCAGTCAATGGTGGGTCAACTGGGGTTACACCA 1059  
Db 784 CAAAGACATACGAGAGCATGAATGGTTTAAACAGGATTTGCCCA 827

## RESULT 10

US-08-557-006C-24  
; Sequence 24, Application US/08557006C  
; Patent No. 6258547  
; GENERAL INFORMATION:  
; APPLICANT: Beri, Rajindar K.  
; APPLICANT: Carling, David  
; APPLICANT: Forster, Robert A.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE  
; FILE REFERENCE: NGAP/PHM37588/UST  
; CURRENT APPLICATION NUMBER: US/08/557,006C



Db 543 ACCGGAGGTCTATCTCAGGAGGCTGTATGCGGGTCTTGAGGTGATATCTGGAGCTGTGG 602  
QY 845 CGTTCCTCTACATCTGTGTGATGCGACCATGCCCTTTGACGGGCGAGGATCAATAAAC 904  
Db 603 TGTATCTCTATGCTCTCTCTGTGTGCGACCCCTCCGTTGCGAGATGAGCACGTCGCTAC 662  
QY 905 ACTGGTGAAGCAATCAGTAAGCGGGCTTACCGGTGAGCGGCCCAAGCCGTCGATGCGCTG 964  
Db 663 GCTCTTTAAGAGATCCGAGGGGTGTGTTCTATCATCCCGAGTATCTCAACCGTTCTAT 722  
QY 965 TGSC---CTGATCCGGTGGCTGTAAATGGTGAACCCACCCGCTGGGCCACACCTGGAGGA 1021  
Db 723 TGCACACTCTGCTGATGCATGCTGCAAGGTGGACCCCTTGAAGCGAGCAACTATCAAGA 782  
QY 1022 TGTAGCCAGTCATGTGGGTCAACTGGGGTTACACCA 1059  
Db 783 CATACGAGAGCATGAATGGTGTAAACAGGATTTGCCCCA 820

## RESULT 12

US-08-677-298-1  
; Sequence 1, Application US/08677298  
; Patent No. 5863729  
; GENERAL INFORMATION:  
; APPLICANT: Pivnica-Worms, Helen  
; TITLE OF INVENTION: DNA SEQUENCES ENCODING HUMAN TCAK-1  
; TITLE OF INVENTION: KINASE  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: USA  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/677,298  
; FILING DATE: 09-JUL-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Caruthers, Jennie M.  
; REGISTRATION NUMBER: 34,464  
; REFERENCE/DOCKET NUMBER: 9-96  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089

; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2698 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 376..2565  
US-08-677-298-1

Query Match 6.2%; Score 179.8; DB 2; Length 2698;  
Best Local Similarity 53.2%; Pred. No. 3.5e-36;  
Matches 404; Conservative 0; Mismatches 352; Indels 3; Gaps 1;  
QY 290 CTACGAGTTCCTGAGACGCTGGCAAGGCGACCTACCGGAAGGTGAAGAGGCGACAGA 349  
Db 540 CTACAGACTGTTGAAACAATCGCAAGGGGAATTTTGCACAAAGTAAATTTGGCAAGACA 599

QY 350 GAGCTCGGGCGTCTGTGGCCATCAAGTCCATCAGSARAGACAAATCAAGATGAGCA 409  
Db 600 TATCTTTAAGGAGAGAGGTTGCAATATAAATAATTTGACAAAACTCAGTGTGAATCCAAC 659  
QY 410 GATCTGCTGCACATACGAGGAGAGATTGAGATCATGTCTTCACTCAACACCCACCCACAT 469  
Db 660 AAGTCTACAAAAGCTCTTACAGAGAGTAAAGATTAATGAAGATTTTAAATCATCCCAATAT 719  
QY 470 CATTCGCATCCATGAAGTGTGTGAGATAGCAGCAAGATTGTGATGTGTCATGAGATGTC 529  
Db 720 AGTGAAGTTATTTCGAAGTCAATGAAACTGAAAAACACTCTACCTATCATGGAATATGC 779  
QY 530 CAGCCGAGGCGATCTGTATGATTACATCAGTGAGCGGCCACCGCTGAGTCAGCGGACGC 589  
Db 780 AAGTGGAGGTGAAGTATTTGACTATTGTTGTCACATGGCAGGATGAAGGAAAAAGAGC 839  
QY 590 CAGGCATTTTCTCCGACAGATCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 649  
Db 840 AAGATCTAAATTTAGACAGATTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 899  
QY 650 TCACCGAGATCTCAAGTGGAAAAACATCCTTTAGATGCCAATGGAACATCAAGATTGC 709  
Db 900 ACATCGAGACCTCAAGGCTGAAAAATCTATTGTTAGATGCCGATATGAACATTTAAATAGC 959  
QY 710 TGACTTTGGCTCTCCAACTGTACACAAAGGCAAGTTCTCCAGACGTTCTGTGGGAG 769  
Db 960 AGATTTGCGTTTGTAGCAATGAATTTACTGTTGGCGGTAAACTCGACACGTTTGTGGCAG 1019  
QY 770 CCCTCTCTACGCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCGAGAGTGA 829  
Db 1020 TCCTCATACGACGACCTGAGCTCTTCCAGGGCAAGAAATATGACGGGCCAGAGTGA 1079  
QY 830 CAGCTGCTCTGCGGGTCTCTCTGTACATCTCTGTGTCATGCGCACCATGCTTGAACG 889  
Db 1080 TGTGTGAGTCTGCGGGTCAATTTTATACACACTAGTCAGTGGGTCACTTCCCTTTGATGG 1139  
QY 890 GCAGGATCATAAACACTGTTGAAGCAAAATCAGTAACGGGGCTTACCGTGAGCGCCCAA 949  
Db 1140 GCAAACTTAAGAACTGAGAGAGAGATTAAGAGGGAATACAGAATTCCTCTTA 1199  
QY 950 GCGTCCGATGCTGTG---GCCTGATCCGGTGGCTGTTAATGGTGAACCCACCCCGTCG 1006  
Db 1200 CATGCTCTACAGACTGTGAAAACTTCTCAAACTGTTCTGCTGCTGCTGCTGCTGCTGCTG 1259  
QY 1007 GGCACACTGGAGGATGTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1045  
Db 1260 CGGCACTCTAGACAAATCATGAAGGACAGGTGGATCAA 1298

## RESULT 13

US-09-774-528-112  
; Sequence 112, Application US/09774528  
; Patent No. 6743619  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Zhang, Jie  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 802  
; CURRENT APPLICATION NUMBER: US/09/774,528

; CURRENT FILING DATE: 2001-01-30  
; NUMBER OF SEQ ID NOS: 441  
; SOFTWARE: pt\_FL\_genes Version 2.0  
; SEQ ID NO 112  
; LENGTH: 2899  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (119)..(2899)  
US-09-774-528-112

Query Match 6.1%; Score 176.4; DB 4; Length 2899;  
Best Local Similarity 52.4%; Pred. No. 2.7e-35;  
Matches 412; Conservative 0; Mismatches 371; Indels 3; Gaps 1;

```
QY 235 TGATGAGAGCAGCGGTGAAGCGGCACCATCACAAACACAACTCTGGCAGCGCTACG 294
DB 120 TGGTCATGGCGGATGGCCCGAGGCACCTTGACGCGCGCGGTCCGGGTGGGTCTACG 179
QY 295 AGTTCTCGAGAGCCTGGGCAAGGGCACCTACGGGAAGGTGAAGAGGCACGAGAGCT 354
DB 180 ACATCGAGGACGCTGGGCAAGGGCACTTCGCTGTGTGAAGCTGGGCGGCACCGGA 239
QY 355 CGGGCGCTCTGGTGCCATCAAGTCCATCAGGAAGA CAAATCAAAAGATGAGCAGGATC 414
DB 240 TCACCAAGACGGAGTGGCAATAAATAATCGATAAGTCTCAGCTGGATGCAATGACC 299
QY 415 TGTGACATACGGAGGAGATGAGATCAATGCTTCACTCAACACCCACATCATTTG 474
DB 300 TTGAGAAATCTACCGAGAGATGACAAATAATGAAATGTTAGACCACTCATATATCA 359
QY 475 CCATCCATGAAGTGTGGAGATAGCAGCAAGATTGTTGTTGTCATGAGATGTCAGGCC 534
DB 360 AACTTTACAGTAATGAGACCAAAAGTATGTTGACCTTGTGACAGATATGCCAAA 419
QY 535 GAGCGCATCTGTATGATTACATCAGTAGGCGGCACCGCTGAGTAGCGGAGCCAGGC 594
DB 420 ATGGAGAAATTTTGAATCTTCTAATCATGCGCGTTTAAATGATCTGAGCCAGGC 479
QY 595 ATTTCTCCAGAGATCGTGTGCGCTGCACTACTGCCACAGAACGGGATCGTTCAAC 654
DB 480 GAAAAATCTCGCAAAATCTCTGCTGTGATTAATGTCATGTTGCGGAAGATTGTGACC 539
QY 655 GAGATCTCAAGCTGGAAAAATCTCTAGATGCAATGGAACATCAAGATTGCTGACT 714
DB 540 GTGACCTCAAGCTGAAAAATCTCTGCTGGATACAAATGATGATCAAAATAGCAGAT 599
QY 715 TTGGCTCTCCAACTGTACCAAAAGGCAAGTTCTCTCCAGAGCTTCTGTGGAGCCCTC 774
DB 600 TCGGTTTGGAAATTTCTTTAAAGTGTGAACTGCTGGCAACATGGTGGCAGCCCC 659
QY 775 TCTACGCTCGCTGAGATGATCAACGGGAAGCCCTATGTGGGCGCAGAGGTGACAGCT 834
DB 660 CTTATGAGCCCGCAGAAAGTCTTTGAAGGCGAGCAGTATGAAGGACCAACAGCTGGA 719
QY 835 GGTCTCTGGCGCTTCTCTGTACATCTGTGTGATGCAATGCAATGCTTTGAGCGGAGG 894
DB 720 GAGATATGGGAGTTGTTCTTATGTCTGTGTGAGAGTGTGCGCTTTGATGGACCGA 779
QY 895 ATCAAAACACCTGGTGAAGCAATCAGTAACCGGGGCTTACCGTAGCCGCCCAAGCCGT 954
DB 780 CTCTTCAATTTTGAAGCAGAGGTTCTGGAGGAAGATTCCGGATTCGATTTTCATGT 839
QY 955 CCGATGCTGTG---GCCTGATCGGTGGCTGTTTAAATGTTGAAACCCCAACCGCGGCCA 1011
DB 840 CAGAAAGATTGAGACACCTTATCCGAAGGATGTTGTTCTTAGACCCATCAAAACGGCTAA 899
QY 1012 CACTGG 1017
DB 900 CCATAG 905
```

## RESULT 14

US-09-930-181-3  
; Sequence 3, Application US/09930181  
; Patent No. 6455292  
; GENERAL INFORMATION:  
; APPLICANT: Origene Technologies  
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas  
; FILE REFERENCE: 16U 101 V1  
; CURRENT APPLICATION NUMBER: US/09/930,181  
; CURRENT FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 3  
; LENGTH: 3364  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (482)..(2239)  
US-09-930-181-3

Query Match 6.0%; Score 174; DB 4; Length 3364;  
Best Local Similarity 53.4%; Pred. No. 1.2e-34;  
Matches 414; Conservative 0; Mismatches 355; Indels 7; Gaps 2;

```
QY 288 CGCTACGAGTTCCTGGAGACGCTGGGCAAGGCACTTACGGGAAGGTGAAGAGGCACGA 347
DB 280 CCCTACCGGCTGGAGAAGACGCTGGGCNAGGGGCAGACAGTCTGGTGAAGCTGGGGTT 339
QY 348 GAGAGCTCGGGGCGTCTGTGGCCATCAAGTCCATCAGGAAGACAAAATCAAAAGATGAG 407
DB 340 CACTGGCTCACCTGCCAGAAGGTGGCCATCAAGATCGTCAACGCTGAGAAGCTCAGCGAG 399
QY 408 CAGGATCTCTGCACATACGGAGGAGATTGAGATCATGCTTCACTCAACACCCCCAC 467
DB 400 TCGGTCTGATGAAGGTGGAGCGGAGATGCGCATCTCTGAAGCTCATTTGAGCACCCCCAC 459
QY 468 ATCATTTGCCATCCATCAAGTGTGAGAAATAGCAGCAAGATTCTGATT----GTCAATGGA 523
DB 460 GTCTAAAGCTGCACAGCTTTATGAANACAAAATATTTGTAGTACCTGGTGTCTAGA 519
QY 524 GTATGCCAGCCGAGGCGATCTGTATGATTACATCAGTAGCGGCCACGCGTGTGAGTACG 583
DB 520 ACAGTGTCAAGTGTGAGCTCTTCGACTACCTGCTGAGGAGGAGGAGGTGACGCGCTAA 579
QY 584 GGAGCCAGGCATTTCTTCGACAGATCTGTCTGCCCTGCACTACTGCCACAGAACCG 643
DB 580 GGAGGCTCGGAAGTTCTTCGCGCAGATCATCTCTGCGCTGGAATCTCTGCCACAGCCACTC 639
QY 644 GATCGTTTCCCGAGATCTCAAGCTGGAATAACATCTTCTAGATGCCAATGGAATAACATCAA 703
DB 640 CATATGCCACAGGATCTGAAACCTGAAACCTCTCTGCTGGACGAGNAGAACATCCG 699
QY 704 GATTGTGATTTTGGCTCTCCAACTGTACCAAAAGGCAAGTTCCTTCAGAGCTTTCTG 763
DB 700 CATCGCAGACTTTGGCATGGCGCTCCCTGCAAGTTGGCGACAGCTGTTGGAGACCACTG 759
QY 764 TGGAGCCCTCTCTAGCCCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCCAGA 823
DB 760 TGGGTCCCCCCTACGCTGCGGAGGTGATCCGGGGGGGAGAAATGATACGCGCCGGA 819
QY 824 GGTGGAGACGTGCTCTGGGCGCTTCTCTGTATCATCTCTGTGATGGCACCACATGCCCTT 883
DB 820 GCGGAGCTGTGAGCTGCGGCGCTCATCTGTTGCGCTTGTCTGTGGGGCTCTGCGCTT 879
QY 884 TGACGGCAGGATCATAAACACTGGTGAAGCAAAATCAGTAAACGGGCTTACCGTGAAGCC 943
DB 880 CGACGATGACAACTTGGCAGCTGTCTGGAGAAGGTGAAGCGGGGCGTGTGTCACATGCC 939
QY 944 GCCC---AAGCCGTCGATGCTGGCTGATCCGGTGGCTGTTTAAATGTTGTAACCCAC 1000
DB 940 GCACCTTTATCCCGCCGACTGCGCAGAGTCTGCTACGGGGGATGATCGAGGTGACGCGCC 999
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 20:17:54 ; Search time 1516.67 Seconds  
(without alignments)  
10994.228 Million cell updates/sec

Title: US-09-980-464-4  
Perfect score: 2902  
Sequence: 1 cactagtggatccaaagaat.....aaaaaaaaaaaaaaaaaaaaa 2902

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2902	100.0	2902	15	US-10-355-975-4
2	2863.4	98.7	3073	17	US-10-322-281-22
3	2066.8	71.2	2929	17	US-10-343-514-1
4	2022.2	69.7	2027	17	US-10-343-514-87
5	1938.8	66.8	37278	17	US-10-322-281-21
6	1771.4	47.5	2026	17	US-10-343-514-27
7	1379	47.5	3353	9	US-09-963-159-1
8	1379	47.5	3353	16	US-10-423-543-43
9	1375.8	47.4	3404	17	US-10-322-281-25
10	1374.2	47.4	3443	18	US-10-370-715B-639
11	1374.2	47.4	3463	18	US-10-618-941-11
12	1372.2	47.3	2501	16	US-10-302-172-215
Sequence 4, Appli					
Sequence 22, Appl					
Sequence 1, Appli					
Sequence 87, Appl					
Sequence 21, Appl					
Sequence 27, Appl					
Sequence 1, Appli					
Sequence 43, Appl					
Sequence 25, Appl					
Sequence 639, Appl					
Sequence 11, Appl					
Sequence 215, Appl					

13	1370.6	47.2	3360	16	US-10-311-034-38	Sequence 38, Appl
14	1360.2	46.9	1884	9	US-09-963-159-3	Sequence 3, Appli
15	1360.2	46.9	1884	16	US-10-423-543-45	Sequence 45, Appl
16	1071.4	36.9	1186	17	US-10-343-514-13	Sequence 13, Appl
17	943.4	32.5	1186	17	US-10-343-514-70	Sequence 70, Appl
18	849.2	29.3	2616	17	US-10-363-829-151	Sequence 151, Appl
19	849.2	29.3	2619	17	US-10-363-829-250	Sequence 250, Appl
20	732	25.2	39699	17	US-10-322-281-24	Sequence 24, Appl
21	675.4	23.3	741	17	US-10-343-514-32	Sequence 32, Appl
22	648	22.3	707	17	US-10-343-514-91	Sequence 91, Appl
23	648	22.3	712	17	US-10-343-514-31	Sequence 31, Appl
24	644.2	22.2	661	17	US-10-343-514-93	Sequence 93, Appl
25	643	22.2	734	17	US-10-343-514-92	Sequence 92, Appl
26	583.2	20.1	588	17	US-10-343-514-97	Sequence 97, Appl
27	557.4	19.2	660	17	US-10-343-514-33	Sequence 33, Appl
28	534	18.4	578	17	US-10-343-514-95	Sequence 95, Appl
29	515	17.7	548	17	US-10-343-514-96	Sequence 96, Appl
30	513	17.7	521	17	US-10-343-514-94	Sequence 94, Appl
31	511.6	17.6	579	17	US-10-343-514-35	Sequence 35, Appl
32	507.6	17.5	6828	9	US-09-780-949-1	Sequence 1, Appli
33	507.6	17.5	6828	15	US-10-354-358-81	Sequence 81, Appl
34	507.6	17.5	6828	15	US-10-172-118-1541	Sequence 1541, Ap
35	507.6	17.5	6828	16	US-10-342-887-1541	Sequence 1541, Ap
36	507.6	17.5	6828	18	US-10-723-860-2552	Sequence 2552, Ap
37	507.6	17.5	6854	18	US-10-723-860-6703	Sequence 6703, Ap
38	506.4	17.5	1986	9	US-09-780-949-3	Sequence 3, Appli
39	506.4	17.5	2884	9	US-09-780-949-5	Sequence 5, Appli
40	493.6	17.0	585	17	US-10-343-514-37	Sequence 37, Appl
41	483.8	16.7	548	17	US-10-343-514-36	Sequence 36, Appl
42	474.6	16.4	521	17	US-10-343-514-34	Sequence 34, Appl
43	383.4	13.2	439	17	US-10-343-514-10	Sequence 10, Appl
44	381.4	13.1	436	17	US-10-343-514-14	Sequence 14, Appl
45	365.6	12.6	421	17	US-10-343-514-9	Sequence 9, Appli

## ALIGNMENTS

## RESULT 1

US-10-355-975-4  
; Sequence 4, Application US/10355975  
; Publication No. US20030162277A1  
; GENERAL INFORMATION:  
; APPLICANT: Immunex Corporation  
; APPLICANT: Bird, Timothy A.  
; APPLICANT: Virca, G. Duke  
; APPLICANT: Martin, Unja  
; APPLICANT: Anderson, Dirk M.  
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES  
; FILE REFERENCE: 2923-A  
; CURRENT APPLICATION NUMBER: US/10/355,975  
; PRIOR FILING DATE: 2003-01-30  
; PRIOR APPLICATION NUMBER: US/09/579,664B  
; PRIOR FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 2902  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-355-975-4

Query Match	100.0%	Score 2902;	DB 15;	Length 2902;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2902;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CAC TAGTGGATCCAAAGAAATTCGGCACGAGGCGTGTCTCGGGTGGCGCTGTGACCTCTGAG	60	
Db	1	CAC TAGTGGATCCAAAGAAATTCGGCACGAGGCGTGTCTCGGGTGGCGCTGTGACCTCTGAG	60	
Qy	61	CCCGGGCTCAGGCGCGCGTGTCTGTGTCGCCGACCCACTCCACCTCGCGGTCCCGCA	120	
Db	61	CCCGGGCTCAGGCGCGCGTGTCTGTGTCGCCGACCCACTCCACCTCGCGGTCCCGCA	120	

Qy	121	CCATGGAGTCGGTGGCCCTTACTCCAGCGCCCGAGCCAGGCTCCCTCGGCTCCGCCCTGG	180
Db	121	CCATGGAGTCGGTGGCCCTTACTCCAGCGCCCGAGCCAGGCTCCCTCGGCTCCGCCCTGG	180
Qy	181	CCTCGGAGAGCCCGCGCGCTGGCGGACGGGCTCATCAAGTCGCTAAACCTCTGATGA	240
Db	181	CCTCGGAGAGCCCGCGCGCTGGCGGACGGGCTCATCAAGTCGCTAAACCTCTGATGA	240
Qy	241	AGAGCAGGGGCTGAAGGGGACCATCAAAACACCAACCTCGGACCGCTACAGATTCC	300
Db	241	AGAGCAGGGGCTGAAGGGGACCATCAAAACACCAACCTCGGACCGCTACAGATTCC	300
Qy	301	TGGAGAGCGCTGGGCAAGGGCACCTACGGGAAGGTGAAGAGGACAGAGAGCTCGGGGC	360
Db	301	TGGAGAGCGCTGGGCAAGGGCACCTACGGGAAGGTGAAGAGGACAGAGAGCTCGGGGC	360
Qy	361	GTCTGGTGGCCATCAAGTCCATCAGGAAGAACAATAACAAAGATGAGCAGGATCTGCTC	420
Db	361	GTCTGGTGGCCATCAAGTCCATCAGGAAGAACAATAACAAAGATGAGCAGGATCTGCTC	420
Qy	421	ACATACGGAGGAGATTGAGATCATCTCTTCACTCAACCCACCCACCATATGTCATCC	480
Db	421	ACATACGGAGGAGATTGAGATCATCTCTTCACTCAACCCACCCACCATATGTCATCC	480
Qy	481	ATGAAGTGTGAGATAGCAGCAAGATTGTGATTGTATGAGTATGCCAGCGAGGCG	540
Db	481	ATGAAGTGTGAGATAGCAGCAAGATTGTGATTGTATGAGTATGCCAGCGAGGCG	540
Qy	541	ATCTGTATGATTAATCATCAGTGAAGCGGACGGCTGAGTGAAGCGGAGCCAGGCAATTTCT	600
Db	541	ATCTGTATGATTAATCATCAGTGAAGCGGACGGCTGAGTGAAGCGGAGCCAGGCAATTTCT	600
Qy	601	TCGACAGATCGTGTCCCTGCACTACTGCTCCACAGACAGGAGCGGATCGTTCACGAGATC	660
Db	601	TCGACAGATCGTGTCCCTGCACTACTGCTCCACAGACAGGAGCGGATCGTTCACGAGATC	660
Qy	661	TCAAGCTTGGAAAAATCCTTCTAGATGCCAATGAAACATCAAGATTGCTGACTTTGGCC	720
Db	661	TCAAGCTTGGAAAAATCCTTCTAGATGCCAATGAAACATCAAGATTGCTGACTTTGGCC	720
Qy	721	TCTCCAACTGTACCAAAAGGCAAGTTCTCCAGACGTTCTGFGAGCGCTCTCTACG	780
Db	721	TCTCCAACTGTACCAAAAGGCAAGTTCTCTCCAGACGTTCTGFGAGCGCTCTCTACG	780
Qy	781	CCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGCCAGAGGTGGAAGCTGTCTC	840
Db	781	CCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGCCAGAGGTGGAAGCTGTCTC	840
Qy	841	TGGGCGTTCTCCTGTATCATCTGTGTGATGGCACCATGCGCTTTGACGGGACAGGATCATA	900
Db	841	TGGGCGTTCTCCTGTATCATCTGTGTGATGGCACCATGCGCTTTGACGGGACAGGATCATA	900
Qy	901	AAACACTGGTGAAGCAAAATCAGTAAACGGGCTTACGTGAGCGGCGCCAAAGCGTCCGATG	960
Db	901	AAACACTGGTGAAGCAAAATCAGTAAACGGGCTTACGTGAGCGGCGCCAAAGCGTCCGATG	960
Qy	961	CCTGTGGCCGTATCCGGTGGCTGTTAATGTGAACCCACCCCTCGGCGCACACTGGAGG	1020
Db	961	CCTGTGGCCGTATCCGGTGGCTGTTAATGTGAACCCACCCCTCGGCGCACACTGGAGG	1020
Qy	1021	ATGTAGCAGTCAATGTTGGTGGTCAACTGGGTTTACACCCAGGAGTCGGGGAACAGGAAG	1080
Db	1021	ATGTAGCAGTCAATGTTGGTGGTCAACTGGGTTTACACCCAGGAGTCGGGGAACAGGAAG	1080
Qy	1081	CCCTGGTGAAGGTGGGACCCCTAGTGTGACTTTTGGCGGGGCTCCATGCGGACTGGT	1140
Db	1081	CCCTGGTGAAGGTGGGACCCCTAGTGTGACTTTTGGCGGGGCTCCATGCGGACTGGT	1140
Qy	1141	TAGTGTCTCCTCGGCGCCCTCTCGGAGAAATGAGCCAGAGGTGTCAGCTTCTCAAGC	1200
Db	1141	TAGTGTCTCCTCGGCGCCCTCTCGGAGAAATGAGCCAGAGGTGTCAGCTTCTCAAGC	1200

Qy	1201	AGCACGTGCGGAGGTGGAGACATGTACTGGCTGGAGCGGCAACATTTCTTTAAGA	1260
Db	1201	AGCACGTGCGGAGGTGGAGACATGTACTGGCTGGAGCGGCAACATTTCTTTAAGA	1260
Qy	1261	AGTCCCGAAAGAGAAATGATAGTGGCTCAAAATCTGAAGGTGACCCGGCTGAGGATACCT	1320
Db	1261	AGTCCCGAAAGAGAAATGATAGTGGCTCAAAATCTGAAGGTGACCCGGCTGAGGATACCT	1320
Qy	1321	CTTCTGCGCTGGCAAGAGCAGCTTAAAGTTCGAAAGGCAATTTCAAGAAAAGTCTC	1380
Db	1321	CTTCTGCGCTGGCAAGAGCAGCTTAAAGTTCGAAAGGCAATTTCAAGAAAAGTCTC	1380
Qy	1381	CTACTCTGTCAGGGAGGTACAGAGGACCTCAGGAATCTCAGACCGGTGCTGATACTC	1440
Db	1381	CTACTCTGTCAGGGAGGTACAGAGGACCTCAGGAATCTCAGACCGGTGCTGATACTC	1440
Qy	1441	CAGGCGAGCCTGTCCCTGTCTGTATCTCTGCTCCCAAGGAAGGCATCTTTAAGAACTCTC	1500
Db	1441	CAGGCGAGCCTGTCCCTGTCTGTATCTCTGCTCCCAAGGAAGGCATCTTTAAGAACTCTC	1500
Qy	1501	GACAGCGTGAATCTGGTTACTACTCTCTCCAGAGCCAGCGAGTCTGGGGAATCTCTTAG	1560
Db	1501	GACAGCGTGAATCTGGTTACTACTCTCTCCAGAGCCAGCGAGTCTGGGGAATCTCTTAG	1560
Qy	1561	ACGCCAGTGTATGTTGTGAGTGGGACCCCGTGGAGCAGAGTCTCCACAGGCTTCAG	1620
Db	1561	ACGCCAGTGTATGTTGTGAGTGGGACCCCGTGGAGCAGAGTCTCCACAGGCTTCAG	1620
Qy	1621	GGCTCTCTCTCCACCGCAAGGGCATTTCTCAAACTCAATGGCAAGTTCTCCCGCACAGCCT	1680
Db	1621	GGCTCTCTCTCCACCGCAAGGGCATTTCTCAAACTCAATGGCAAGTTCTCCCGCACAGCCT	1680
Qy	1681	TAGAAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCACTGGGCTCTCTCCATCTTG	1740
Db	1681	TAGAAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCACTGGGCTCTCTCCATCTTG	1740
Qy	1741	CAGCCCGGCCAGCGCCCTCAGGGGCTGAGTGAGGACAGCATCTCTCCCGAGT	1800
Db	1741	CAGCCCGGCCAGCGCCCTCAGGGGCTGAGTGAGGACAGCATCTCTCCCGAGT	1800
Qy	1801	CCTTTGACCAATTTGAGCTTGCCTGAAACGTTCTCCGAAACCCCACTGAGGGGCTGTGTGT	1860
Db	1801	CCTTTGACCAATTTGAGCTTGCCTGAAACGTTCTCCGAAACCCCACTGAGGGGCTGTGTGT	1860
Qy	1861	CTGTGAGCAACTGAGGGGCTTTGAGCGCTCCCTCAGAAAGTCTGAAGCATGTGTGC	1920
Db	1861	CTGTGAGCAACTGAGGGGCTTTGAGCGCTCCCTCAGAAAGTCTGAAGCATGTGTGC	1920
Qy	1921	AGGAATCTTTGGGGATAGTCTGCTTCTGACAGACTGCCAAGAGGTGACTGCAAGCCT	1980
Db	1921	AGGAATCTTTGGGGATAGTCTGCTTCTGACAGACTGCCAAGAGGTGACTGCAAGCCT	1980
Qy	1981	ACAGACAAGCCCTAGGAATCTGCTCAAAAGTCTCAGCTGAGGAAGGAGATGGTCCCTAGT	2040
Db	1981	ACAGACAAGCCCTAGGAATCTGCTCAAAAGTCTCAGCTGAGGAAGGAGATGGTCCCTAGT	2040
Qy	2041	ATGGGGTGGCTCTGAGGGGTTTTCAGAGGAACTTCCAGAGGATCTCCAGTGAATA	2100
Db	2041	ATGGGGTGGCTCTGAGGGGTTTTCAGAGGAACTTCCAGAGGATCTCCAGTGAATA	2100
Qy	2101	GAGTACATCAAGGGCTCTACGCTGCGAGCTGAACTCAAAAGATGAGAGAAATCGC	2160
Db	2101	GAGTACATCAAGGGCTCTACGCTGCGAGCTGAACTCAAAAGATGAGAGAAATCGC	2160
Qy	2161	ATTGATGTGGAAGGAATGGGAACTTGTCTCCCGAGTGTATAGTGGGGTGGCTGAA	2220
Db	2161	ATTGATGTGGAAGGAATGGGAACTTGTCTCCCGAGTGTATAGTGGGGTGGCTGAA	2220
Qy	2221	GGTGGCTACTCTTTGTGGCCATGAGTGTACCCATGACATTTCCACCTCTCTCGG	2280
Db	2221	GGTGGCTACTCTTTGTGGCCATGAGTGTACCCATGACATTTCCACCTCTCTCGG	2280
Qy	2281	CTGCACTTTCATTAAGTTTCTGTTTCCATCAACCCAGGGGTAGAAACCTTGACTTCT	2340

```
Db 2281 |||||CTGCACCTTACATAAGTTTCTGTTTCCATCAACACCAGGGTTAGAACCCCTGACTTCT 2340
Qy 2341 GGGAGGTAATGTGTAGTGAAGTGCCTTATTTAGAGAGAAACAGCCCTCTGTTTCCATCT 2400
Db 2341 GGGAGGTAATGTGTAGTGAAGTGCCTTATTTAGAGAGAAACAGCCCTCTGTTTCCATCT 2400
Qy 2401 CTGCTGTGTGCATCTCAAAAGACCTGGGAAAGACTCGGACCGCTGTTTGACTTCAATCTCAA 2460
Db 2401 CTGCTGTGTGCATCTCAAAAGACCTGGGAAAGACTCGGACCGCTGTTTGACTTCAATCTCAA 2460
Qy 2461 GGGGACAGATGCCCTGGACCCCATCTTAGATCTCAGAGACTTGAACCTTGAAGCTGTT 2520
Db 2461 GGGGACAGATGCCCTGGACCCCATCTTAGATCTCAGAGACTTGAACCTTGAAGCTGTT 2520
Qy 2521 CCTAGTACCCAGATGTGGATGGATGCTCTGTTCTCAGGCAACCGGACCTAGAAATGTC 2580
Db 2521 CCTAGTACCCAGATGTGGATGGATGCTCTGTTCTCAGGCAACCGGACCTAGAAATGTC 2580
Qy 2581 TGACTTATTTATTTTGTGATTTCTCACTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 2640
Db 2581 TGACTTATTTATTTTGTGATTTCTCACTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 2640
Qy 2641 TTTTGTGTTTAAAGTGAATTTGCTGCTTTCATATGTAATGTAATGTAATGTAATGTAATGTA 2700
Db 2641 TTTTGTGTTTAAAGTGAATTTGCTGCTTTCATATGTAATGTAATGTAATGTAATGTAATGTA 2700
Qy 2701 TCCACTGTGCACTGAAGTTTATGTACAGAGAGATTTTGGCAATGATGTCCTCTATTTTC 2760
Db 2701 TCCACTGTGCACTGAAGTTTATGTACAGAGAGATTTTGGCAATGATGTCCTCTATTTTC 2760
Qy 2761 AAGGGGGGTGGGGCGGTTTTTCAATGTATGTCTTGAGCACTGTCTGGATTGAGTCTCCA 2820
Db 2761 AAGGGGGGTGGGGCGGTTTTTCAATGTATGTCTTGAGCACTGTCTGGATTGAGTCTCCA 2820
Qy 2821 GTCCCTTACACCAAGGCTGGGCACCCCTCCCTCATCTTCATCTGTGGCCAAAAA 2880
Db 2821 GTCCCTTACACCAAGGCTGGGCACCCCTCCCTCATCTTCATCTGTGGCCAAAAA 2880
Qy 2881 AAAAAAAAAAAAAAAAAAAAAA 2902
Db 2881 AAAAAAAAAAAAAAAAAAAAAA 2902
```

## RESULT 2

```
US-10-322-281-22
; Sequence 22, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 52945201000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 3073
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-322-281-22
```

```
Query Match 98.7%; Score 2863.4; DB 17; Length 3073;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2864; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 32 CGTGCTCGGGTGGGCTGTGACCTCTGAGCCCGGCTCAGCGCGCTGCTACTGCTGC 91
Db 1 CGTGCTCGGGTGGGCTGTGACCTCTGAGCCCGGCTCAGCGCGCTGCTACTGCTGC 60
Qy 92 CCGACCCACTCCACTCGCGGTCCCGCACCATGAGAGTGGGTGGCCCTTACTCCAGCGCCC 151
Db 151 CCGACCCACTCCACTCGCGGTCCCGCACCATGAGAGTGGGTGGCCCTTACTCCAGCGCCC 1200
```

```
Db 61 CCGACCCACTCCACTCGCGGTCCCGCACCATGAGAGTGGGTGGCCCTTACTCCAGCGCCC 120
Qy 152 GAGCCAGGCTCCCTCGGCTCCGCTCCGCTCGGAGAGCGCCGCGCTGGCGGACGG 211
Db 121 GAGCCAGGCTCCCTCGGCTCCGCTCCGCTCGGAGAGCGCCGCTGGCGGACGG 180
Qy 212 GCTCATCAAGTCCGCTAAACCTCTGATGAAGAGAGCGGCTGAAGCGGACCATCAAA 271
Db 181 GCTCATCAAGTCCGCTAAACCTCTGATGAAGAGAGCGGCTGAAGCGGACCATCAAA 240
Qy 272 ACACAACTCGGCGACCGCTACAGTTCCTGGAGAGAGCTGGGCAAGGCAACCTACGGGAA 331
Db 241 ACACAACTCGGCGACCGCTACAGTTCCTGGAGAGAGCTGGGCAAGGCAACCTACGGGAA 300
Qy 332 GGTGAAGAGGACAGAGAGAGCTCGGGCGCTCTGTTGGCCATCAAGTTCATCAGGAAGA 391
Db 301 GGTGAAGAGGACAGAGAGAGCTCGGGCGCTCTGTTGGCCATCAAGTTCATCAGGAAGA 360
Qy 392 CAAAATCAAAGATGAGCAGGATCTGTGCAATACGAGGAGAGATTGAGATCATGTCTTC 451
Db 361 CAAAATCAAAGATGAGCAGGATCTGTGCAATACGAGGAGAGATTGAGATCATGTCTTC 420
Qy 452 ACTCAACACCCCAATCATTTGCCATCCATGAAGTGTGTTGAGAAATAGAGCAAGATTGT 511
Db 421 ACTCAACACCCCAATCATTTGCCATCCATGAAGTGTGTTGAGAAATAGAGCAAGATTGT 480
Qy 512 GATTGTCTGAGGATGATGCCAGCGGAGCGATCTGTATGATTATCATCAGTCAGCGGCCACG 571
Db 481 GATTGTCTGAGGATGATGCCAGCGGAGCGATCTGTATGATTATCATCAGTCAGCGGCCACG 540
Qy 572 GCTGAGTGAAGCGGAGCGCCAGGCAATTTTCCGACAGATCGTGTCTGCCCTGCACTACTG 631
Db 541 GCTGAGTGAAGCGGAGCGCCAGGCAATTTTCCGACAGATCGTGTCTGCCCTGCACTACTG 600
Qy 632 CACACAGAACGGGATCGTTTCAACCGAGATCTCAAGCTGGAAAAATCTCTTAGATGCCAA 691
Db 601 CACACAGAACGGGATCGTTTCAACCGAGATCTCAAGCTGGAAAAATCTCTTAGATGCCAA 660
Qy 692 TGAAGAACATCAAGATTGCTGACTTTGGCCCTCTCCAACTGTACCAAAAGGCAAGTTCCT 751
Db 661 TGAAGAACATCAAGATTGCTGACTTTGGCCCTCTCCAACTGTACCAAAAGGCAAGTTCCT 720
Qy 752 CACAGAGCTTCTGTGGAGCGCTCTCTACGCTCGCTCGCTGAGATAGTCAACCGGAGCGCTA 811
Db 721 CACAGAGCTTCTGTGGAGCGCTCTCTACGCTCGCTCGCTGAGATAGTCAACCGGAGCGCTA 780
Qy 812 TGTGGGCCACAGAGTGGACAGTGTGCTCTGGCGGTTCTCTGTATACATCTCTGTGCAATGG 871
Db 781 TGTGGGCCACAGAGTGGACAGTGTGCTCTGGCGGTTCTCTGTATACATCTCTGTGCAATGG 840
Qy 872 CACCATGCCCTTTGACGGGACGATCATAAACACTGTTGAAAGCAAAATCAGTAACGGGGC 931
Db 841 CACCATGCCCTTTGACGGGACGATCATAAACACTGTTGAAAGCAAAATCAGTAACGGGGC 900
Qy 932 TTACCGTGAAGCGCCCAAGCGCTCCGATGCTGTGGCTGATCCGGTGGCTGTTTAAATGTT 991
Db 901 TTACCGTGAAGCGCCCAAGCGCTCCGATGCTGTGGCTGATCCGGTGGCTGTTTAAATGTT 960
Qy 992 GAACCCCAACCGTGGGCGCACACTGAGAGATGATGACAGTCTGTTGGTGGTCACTGGGG 1051
Db 961 GAACCCCAACCGTGGGCGCACACTGAGAGATGATGACAGTCTGTTGGTGGTCACTGGGG 1020
Qy 1052 TTACACACCGGAGTGGGGGACAGGAAGCCCTCGCTGAGGGTGGGCAACCTAGTGGTGA 1111
Db 1021 TTACACACCGGAGTGGGGGACAGGAAGCCCTCGCTGAGGGTGGGCAACCTAGTGGTGA 1080
Qy 1112 CTTTGGCCGGGCTCCATGCGGACCTGTTTACGTCGCTCTCTCGCGCCCTCTCTCGGAGAA 1171
Db 1081 CTTTGGCCGGGCTCCATGCGGACCTGTTTACGTCGCTCTCTCGCGCCCTCTCTCGGAGAA 1140
Qy 1172 TGAGGCAAGGTTGAGCTTTTCAAGCAGACAGTTCGGGAGGTGGAAAGCACTGTACC 1231
Db 1141 TGAGGCAAGGTTGAGCTTTTCAAGCAGACAGTTCGGGAGGTGGAAAGCACTGTACC 1200
```

```
QY 1232 TGGCTGGAGCGGCAACATTTCTTTAAGAGTCCGGAAGGAGATGACATGGCTCAAAA 1291
Db 1201 TGGCTGGAGCGGCAACATTTCTTTAAGAGTCCGGAAGGAGATGACATGGCTCAAAA 1260
QY 1292 TCTGCAAGGTGACCCGGCTGAGGATACCTCTTTCTGCGCTGGCAAGAGAGCCCTTAAGCT 1351
Db 1261 TCTGCAAGGTGACCCGGCTGAGGATACCTCTTTCTGCGCTGGCAAGAGAGCCCTTAAGCT 1320
QY 1352 TCCGAAAGGCATTTCTCAAGAAAAAGTCTCTTACCTCTGTCAGGGGAGGTACAGGAGGACCC 1411
Db 1321 TCCGAAAGGCATTTCTCAAGAAAAAGTCTCTTACCTCTGTCAGGGGAGGTACAGGAGGACCC 1380
QY 1412 TCAGGAATCAGACCGGTGCTGTATCTCCAGGCGAGCTGCTCCCTGCTATCCCTGCT 1471
Db 1381 TCAGGAATCAGACCGGTGCTGTATCTCCAGGCGAGCTGCTCCCTGCTATCCCTGCT 1440
QY 1472 CCCAAGGAAAGGCATCTTAAAGAGTCTCGACAGCGTGAATCTGGTTACTACTCTCTCC 1531
Db 1441 CCCAAGGAAAGGCATCTTAAAGAGTCTCGACAGCGTGAATCTGGTTACTACTCTCTCC 1500
QY 1532 AGAGCCAGGAGTCTGGGAACTCTTTAGAGCGCAGTGATGTGTTGTGAGTGGGACCC 1591
Db 1501 AGAGCCAGGAGTCTGGGAACTCTTTAGAGCGCAGTGATGTGTTGTGAGTGGGACCC 1560
QY 1592 CGTGGAGCAGAGTCTCCAGAGGCTTCAGGCTCTCTCCACCGCAAGGCAATCTCAA 1651
Db 1561 CGTGGAGCAGAGTCTCCAGAGGCTTCAGGCTCTCTCCACCGCAAGGCAATCTCAA 1620
QY 1652 ACTCAATGGCAAGTCTCCCGCAGAGCTTTAGAGGCACTACCCCTAGCAGCTTTGGCTC 1711
Db 1621 ACTCAATGGCAAGTCTCCCGCAGAGCTTTAGAGGCACTACCCCTAGCAGCTTTGGCTC 1680
QY 1712 CTTGGACCAACTGGCTCTCCCATCTGAGCGCCGCGCCAGCGCCCTCAGGGGCTGT 1771
Db 1681 CTTGGACCAACTGGCTCTCCCATCTGAGCGCCGCGCCAGCGCCCTCAGGGGCTGT 1740
QY 1772 GAGTGGAGCAGCATCTGCTCCGAGTCTTTGACCAATTCGACTTGGCTGACGTCT 1831
Db 1741 GAGTGGAGCAGCATCTGCTCCGAGTCTTTGACCAATTCGACTTGGCTGACGTCT 1800
QY 1832 TCCGGAACCCCACTGAGGGGCTGTGTCTGTGGACAACTGAGGGGGCTTGAGCAGCC 1891
Db 1801 TCCGGAACCCCACTGAGGGGCTGTGTCTGTGGACAACTGAGGGGGCTTGAGCAGCC 1860
QY 1892 TCCTCAGAAAGTCTGAAGCGATGGTGGCAGGAATCTTGGGGATAGCTGCTTTCTCT 1951
Db 1861 TCCTCAGAAAGTCTGAAGCGATGGTGGCAGGAATCTTGGGGATAGCTGCTTTCTCT 1920
QY 1952 GACAGACTGCCAAGAGTGACTGAGCCTACAGAGCCCTAGGATCTGCTCAAGCT 2011
Db 1921 GACAGACTGCCAAGAGTGACTGAGCCTACAGAGCCCTAGGATCTGCTCAAGCT 1980
QY 2012 CAGCTGAGGAAGGAGATGGTGGCTAGTATGGGGTAGGCTCTCAGAGGGTTTGCAGAG 2071
Db 1981 CAGCTGAGGAAGGAGATGGTGGCTAGTATGGGGTAGGCTCTCAGAGGGTTTGCAGAG 2040
QY 2072 AACCTGGGTGCGATTCCTCAGTGAATAGAGTACATCAAGGGCTCTAAGTCTGAGCCT 2131
Db 2041 AACCTGGGTGCGATTCCTCAGTGAATAGAGTACATCAAGGGCTCTAAGTCTGAGCCT 2100
QY 2132 GACTGAACCTGAAAGATGAGAGAAATCGCATTTGATGGAAAGAAATGGGAACCTTGCT 2191
Db 2101 GACTGAACCTGAAAGATGAGAGAAATCGCATTTGATGGAAAGAAATGGGAACCTTGCT 2160
QY 2192 GCCGAGTGTATAGTGGGTGGCTGAGGCTGACCTCTCTTTGTGCCATGAGTGA 2251
Db 2161 GCCGAGTGTATAGTGGGTGGCTGAGGCTGACCTCTCTTTGTGCCATGAGTGA 2220
QY 2252 CCCATGACATTTCCACCCCTGTTCTCTGGCTGACCTTACATAAGTTCTGTTTCATC 2311
Db 2221 CCCATGACATTTCCACCCCTGTTCTCTGGCTGACCTTACATAAGTTCTGTTTCATC 2280
```

```
QY 2312 AACCCACCGGTTAGAACCCCTGACTTCTCTGGGAGGTAATGTGTAGTGAAGTCCCATTTT 2371
Db 2281 AACCCACCGGTTAGAACCCCTGACTTCTCTGGGAGGTAATGTGTAGTGAAGTCCCATTTT 2340
QY 2372 AGAGAGGAACACGCTCTGTTTCCATCTCTGCTGCTGTGTCATCTCAAGAGACCTGGGAAG 2431
Db 2341 AGAGAGGAACACGCTCTGTTTCCATCTCTGCTGCTGTGTCATCTCAAGAGACCTGGGAAG 2400
QY 2432 ACTCGGACCGCTGTTTGAATCTCAAGGGGACAGATGCCCTCGGACCCCATCTTAG 2491
Db 2401 ACTCGGACCGCTGTTTGAATCTCAAGGGGACAGATGCCCTCGGACCCCATCTTAG 2460
QY 2492 ATCTCAGAGACTTGAACCTTGAAGCTGTTCTTAGTACCCAGATGTGGATGGATGCTCTGT 2551
Db 2461 ATCTCAGAGACTTGAACCTTGAAGCTGTTCTTAGTACCCAGATGTGGATGGATGCTCTGT 2520
QY 2552 TTCTCAGGCGCAAGCGGACCTAGAAATGCTGACTTATTTATTTTGTGATTTCTCACTT 2611
Db 2521 TTCTCAGGCGCAAGCGGACCTAGAAATGCTGACTTATTTATTTTGTGATTTCTCACTT 2580
QY 2612 CTGTTTTTGGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG 2671
Db 2581 CTGTTTTTGGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG 2640
QY 2672 AATAATGTGAATGCTGTTCTGGGAACTCCACTGTCGCACTGAAGTTTATGTACAGAG 2731
Db 2641 AATAATGTGAATGCTGTTCTGGGAACTCCACTGTCGCACTGAAGTTTATGTACAGAG 2700
QY 2732 AAGTATTTGGCAATGATGTCCTCTATTCAAGGGGGGTGGGGCGTTTTTCAAAATGTATG 2791
Db 2701 AAGTATTTGGCAATGATGTCCTCTATTCAAGGGGGGTGGGGCGTTTTTCAAAATGTATG 2760
QY 2792 TCTTGAGCACTGTCTGGATGTAGTCTCCAGTCCCTTCAACCAAGGCTGGCCACCTCC 2851
Db 2761 TCTTGAGCACTGTCTGGATGTAGTCTCCAGTCCCTTCAACCAAGGCTGGCCACCTCC 2820
QY 2852 CTCATCTTCATCTGTGGCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2896
Db 2821 CTCATCTTCATCTGTGGCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2865
```

## RESULT 3

```
US-10-343-514-1
; Sequence 1, Application US/10343514
; Publication No. US20040132025A1
; GENERAL INFORMATION:
; APPLICANT: DRUCKER, Daniel J.
; APPLICANT: ROSEN, Cheryl F.
; APPLICANT: LEFEBVRE, Diana L.
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK
; FILE REFERENCE: DPA-DRUC2/PCT
; CURRENT APPLICATION NUMBER: US/10/343,514
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: PCT/CA01/01109
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/222,650
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/274,613
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: CA 2,340,780
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2929
; TYPE: DNA
; ORGANISM: RAT
US-10-343-514-1
```

Query Match 71.2%; Score 2066.8; DB 17; Length 2929;  
Best Local Similarity 86.4%; Pred. No. 0;  
Matches 2488; Conservative 0; Mismatches 277; Indels 115; Gaps 14;

QY 49 GTGACCTCTGAGCGCCGGCTCAGCGCGCTGCTACTGCTGCCCGACCCACCTCCACCTC 108  
Db 9 GTGACCTCTGAGCTCGGCTCTCCGCGCTGCTGCTGCTGCCCGACCCCTCCGCGCTC 68  
QY 109 GCGGTCCCGCCACCATGGAGTGGTGGCTTACTTCAGCGCCCGAGCGCTCCCTCGG 168  
Db 69 GCGGTCCCGCCACCATGGAGTGGTGGCTTACTTCAGCGCCCGAGCGCTCCCTCGG 128  
QY 169 CTTCCGCGCTTGGCTCGGAGAGCGCCGCGCTGGCGGACGGCTCATCAAGTCGGCTA 228  
Db 129 CTTCCGCGCTTGGCACGAGAGCGCCGCGCTGGCGGACGGCTCATCAAGTCGGCTA 188  
QY 229 AACCTCTGATGAAGACGAGCGGTGAAGCGGACCATCAAAACAACCTCGGCGACC 288  
Db 189 AACCTCTGATGAAGACGAGCGGTGAAGCGGACCATCAAAACAACCTCGGCGACC 248  
QY 289 GCTACGAGTTCCTGGAGACCTGGGCAAGGCGACCTACGGGAAGGTGAAGAAAGCACGAG 348  
Db 249 GCTACGAGTTCCTGGAGACCTGGGCAAGGCGACCTACGGGAAGGTGAAGAAAGCACGAG 308  
QY 349 AGAGCTCGGCGCTCTGGTGGCCATCAAGTCCATCAGGAAGACAAAATCAAGATGAGC 408  
Db 309 AGAGCTCGGCGCTCTGGTGGCCATCAAGTCCATCAGGAAGACAAAATCAAGATGAGC 368  
QY 409 AGGATCTGCTGCACATACGAGGAGGATTCAGATCATGTCTTCACTCAACACCCCCACA 468  
Db 369 AGGATCTGTTGCACATCAAGGAGGAGATCGAGATCATGTCTTCACTCAACACCCCCACA 428  
QY 469 TCATTGCCATCCATGAAGTGTGAGAAATAGCAGCAAGATGATGATGATGATGATGATG 528  
Db 429 TCATTGCCATCCATGAAGTGTGAGAAATAGCAGCAAGATGATGATGATGATGATGATG 488  
QY 529 CCAGCCGAGCGGATCTGTATGATTTACATCAGTGAAGCGCCACCGCTGAGTGAGCGGACG 588  
Db 489 CCAGCCGAGCGGATCTGTATGATTTACATCAGTGAAGCGCCACCGCTGAGTGAGCGGACG 548  
QY 589 CCAGGCAATTTCTCCGACAGATCGTGTCTCCCTGCACTACTGCCACAGAGCGGATCG 648  
Db 549 CCAGGCAATTTCTCCGACAGATCGTGTCTCCCTGCACTACTGCCACAGAGCGGATCG 608  
QY 649 TTCCAGGAGATCTCAAGCTGGAACATCTCTTAGATGCAATGGAACATCAAGATTG 708  
Db 609 TTCCAGGAGATCTCAAGCTGGAACATCTCTTAGATGCAATGGAACATCAAGATTG 668  
QY 709 CTGACTTTGGCTCTCCAACTGTACCAAGGCAAGTTCTTCCAGACGTTCTGTGGGA 768  
Db 669 CTGATTTGGCTCTCCAACTGTATCACAAGGCAAGTTCTTCCAGACGTTCTGTGGGA 728  
QY 769 GCCCTCTCTACGCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCGAGAGTGG 828  
Db 729 GCCCTCTCTATGCTCACCTGAGATCGTCAACGGGAAGCCCTATGTGGGCCCGAGAGTGG 788  
QY 829 ACAGCTGGTCTCTGGGGTTCCTGTATCATCTGTGTGATGGACCATCGCCCTTTGACG 888  
Db 789 ACAGCTGGTCTCTGGGGTTCCTGTATCATCTGTGTGATGGACCATCGCCCTTTGACG 848  
QY 889 GGCAGAGATCAATAAACAACCTGTGAAGCAATCAGTAACGGGCTTACCGTGAGCGCCCA 948  
Db 849 GGCAGAGATCAATAAACAACCTGTGAAGCAATCAGTAACGGGCTTACCGTGAGCGCGTGA 908  
QY 949 AGCCGTCCGATGCTGTGGCTGATCCGGTGGCTGTTAAATGGTGAACCCACCCCGTGGG 1008  
Db 909 AACCGTCTGATGCTGTGGCTGATCCGGTGGCTGTTAAATGGTGAATCCCATCCGTCGGG 968  
QY 1009 CCACATGGAGGATGATGACGAGTCAATGGTGGGTCAATCGGGGTTCACACCCCGAGTGC 1068  
Db 969 CCACATGGAGGATGATGACGAGTCAATGGTGGGTCAATCGGGGTTCACACCCCGAGTGC 1028  
QY 1069 GGGAAACAGGAAGCCCTCGTGAAGTGGGACCCCTAGTGTGACTTTGGCCGGGCGCTCCA 1128  
Db 1029 GGGAAACAGGAAGCTCTCGAGAGGGTGGGACCCCTAGCGGTGACTCTTGGCCGGGCGCTCA 1088  
QY 1129 TGGCGGACTGTTACGTGCTCTCTCGCGCCCTCTCTGGAGAATGGAGCAAGGTGTGCA 1188

Db 1089 TGGGAGCTGTTTACGTGCTCTCCGCGCCCTCTCTGGAGAATGGAGCAAAAGTGTGA 1148  
QY 1189 GCTTCTTCAAGCAGCAGCTGCGGAGGTGAAGACACTGTACTTGGGCTGAGCGGCAAC 1248  
Db 1149 GCTTCTTCAAGCAGCAGTGTCCGGAGGTGAAGACCGGGCTGAGCGGCAAC 1208  
QY 1249 ATTCTCTTAAGAAAGTCCGAAAGAGAGATGACATGGCTCAAAATCTCAAGAGTGCACCGG 1308  
Db 1209 ATTCTCTTAAGAAAGTCCGCAAGAGAGATGACATGGCTCAGACTCTCAGAAATGACCCAG 1268  
QY 1309 CTGAGGATACCTTCTCGCCCTGGCAAGAGCAGCCTTAAGCTTCGAAAGGCAATTTCTCA 1368  
Db 1269 TTGAAGATATCTCTCTCGCCCTGGCAAGAAACAGCCTCAAGCTTCGAAAGGTATCTCA 1328  
QY 1369 AGAAAGAGTCTCTACTCTGTCAGGAGGAGTACAGGAGGACCTCAGAACTCAGACCGG 1428  
Db 1329 AGAAAGAGTCTCTCTCTCATCGGGGAGGTACAGGAGGCGCTCAGGAACTCAGACCG 1388  
QY 1429 TGCTGATATCTCCAGGCGAGCTGTCTCTGCTGTATCCCTGCTCCCAAGGAAAGGCAATCC 1488  
Db 1389 TGTCCAATATACCCAGGCGAGCTGTCTCTGCTATACCCCTGCTCCCAAGGAGGCAATTC 1448  
QY 1489 TTAAGAAAGTCTCGACAGCGTGAATCTGGTTACTCTCTCTCCAGAGCCCGTGAGAGTCTG 1548  
Db 1449 TTAAGAAAGTCTCGGACAGCGTGAATCTGGTTACTCTCTCTCCAGAGCCCGAGTGTG 1508  
QY 1549 GGGAACTCTTAGACCGCAGTGTGTTGTGAGTGGGAGCCCGTGAGAGAGTCTC 1608  
Db 1509 GGGAACTCTTAGACCGCAGTGTGTTGTGAGTGGGAGCCCGTGAGAGAGTCTC 1568  
QY 1609 CACAGGCTTCAGGGCTCTCTCCACCGCAAGGCAATCTCAAACTCAATGGCAAGTTCT 1668  
Db 1569 CACAGGCTTCAGGG---GGCTCCATCGCAAGGCAATCTCAAACTCAATGGCAAGTTCT 1625  
QY 1669 CCGCAAGCTTTAGAGGCACTACCCCTAGCAGCTTTGGCTCCCTGGACCAACTGSCCT 1728  
Db 1626 CCGCAAGCTTTAGAGGCACTGCCCCCTAGCAGCTTTGGCTCCCTAGACCAACTGSCCT 1685  
QY 1729 CTTCCCATCTGACAGCCCGCCAGCGCCCTCAGGGGCTGTGAGTGAAGACAGCATCC 1788  
Db 1686 CCGCTCATCTACAGCCCGGCGAGCGCTCCCTCGGAGCTGTGAGTGAAGACAGCATCC 1745  
QY 1789 TGTCTCTCCAGTCTTTGACCAATTTGACTTGTGCTGAACCTTTCCGAAACCCCACTGA 1848  
Db 1746 TGTCTCTCCAGTCTTTGACCAATTTGACTTGTGCTGCGAGCGGCTTCCGAAACCCCACTGA 1805  
QY 1849 GGGGCTGTGTCTGTGGAACAACCTGAGGGGGCTTGAAGAGCCCTCCCTCAGAAAGTCTGA 1908  
Db 1806 GGAGCTGTGTCTGTGGAACAACCTGAGGAGGCTTGAAGAGCCCTCCCTCAGAAAGGCTAA 1865  
QY 1909 AGCGATGGTGGCAGGAATCTTGGGGGATAGCTGCTTTCTCTGACAGACTGCCAAGAGG 1968  
Db 1866 AACGATGGTGGCAGGAATCTTGGGGGATAGCGCGCTTTCTCTGACAGACTGCCAAGAGG 1925  
QY 1969 TGACTGAGCTACAGCAAGCCCTAGGAATCTGCTCAAAAGCTCAGCTGAGGAGGAGAGA 2028  
Db 1926 TGACAGAGGCTACAGCAAGCCCTAGGAATCTGCTGAAAGCTCAGCTGAGGAGGAGAGG 1985  
QY 2029 TGGTGCCCTAGT-ATGGGGTAGGCTCTGAGAGGGTTTGAAGAGAACCCCTGGGTGGGATT 2087  
Db 1986 CAGTGCCCGAGTATGGGGTAGACTCTTAGAGGGGTTTGAAGAGAACCCCTGGGT-AGATT 2044  
QY 2088 CTTCCAGTGAATAGATACATCAAG---GGCTCTAGTCTGAGAGCTGAGCTGAGCTGA 2143  
Db 2045 CCGCAGGGTTGTAGATACATCAAGAACTCTCTCTGCTCTGCTCAGCGCTGATTGAACCTGG 2104  
QY 2144 AAGTGAAGAAATCGCATTTGATGGAAGGAATGGGAACCCCTGTGCTGCCAGTGTGA 2203  
Db 2105 AGCTGAGAGAAATAGCAGAGATATGGAAGGACTGACCTAGAGTCTGACTGACAGATG 2164  
QY 2204 TAGTGGGGT---GGCCTGAAGGTGCTCTCTCTTTGTGCTGCTGAGTGTCAACCATGACA 2260





```
QY 1069 GGGACACAGGAAGCCCTGCTGAGGGTGGGACACCTAGTGTGACTTTGGCCGGGCTCTCA 1128
Db 1021 GGGACACAGGAAGCCCTGCTGAGGGTGGGACACCTAGTGTGACTTTGGCCGGGCTCTCA 1080
QY 1129 TGGCGGACTGCTTACGTGCTCTCTGGGCCCTCTCTGGAGATGAGACCAAGGTGTGA 1188
Db 1081 TGGCGGACTGCTTACGTGCTCTCTGGGCCCTCTCTGGAGATGAGACCAAGGTGTGA 1140
QY 1189 GCTTCTTCAAGCAGCAGCTGCCGGAGGTGGAAGCAGTGTACCTGGCTGGCTGGAGCGGCAAC 1248
Db 1141 GCTTCTTCAAGCAGCAGCTGCCGGAGGTGGAAGCAGTGTACCTGGCTGGAGCGGCAAC 1200
QY 1249 ATTCTCTTAAGAAAGTCCGAAAGGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGG 1308
Db 1201 ATTCTCTTAAGAAAGTCCGAAAGGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGG 1260
QY 1309 CTGAGGATACCTCTTCTGCGCTTGGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATCTCA 1368
Db 1261 CTGAGGATACCTCTTCTGCGCTTGGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATCTCA 1320
QY 1369 AGAAAGAGTCTCTACCTCTCAGGGAGGTACAGGAGGACCTCAGGAACTCAGACCGG 1428
Db 1321 AGAAAGAGTCTCTACCTCTCAGGGAGGTACAGGAGGACCTCAGGAACTCAGACCGG 1380
QY 1429 TGCTGTATCTCCAGGGCAGCTGTCTCTGTATCTCTGCTCCCAAGGAAAGGCAATCC 1488
Db 1381 TGCTGTATCTCCAGGGCAGCTGTCTCTGTATCTCTGCTCCCAAGGAAAGGCAATCC 1440
QY 1489 TTAAGAAAGTCTGACAGCAGTGAATCTGGTTACTACTCTCTCAGAGCCAGCAGTCTG 1548
Db 1441 TTAAGAAAGTCTGACAGCAGTGAATCTGGTTACTACTCTCTCAGAGCCAGCAGTCTG 1500
QY 1549 GGGAACTCTTAGAGCCAGTGTGTGTGAGTGGGACCCCGTGGAGCAGAGTCTC 1608
Db 1501 GGGAACTCTTAGAGCCAGTGTGTGTGAGTGGGACCCCGTGGAGCAGAGTCTC 1560
QY 1609 CACAGGCTTCAGGGCTCTCTCTCACCGCAAGGGCAATTTCAAACTCAATGGCAAGTTCT 1668
Db 1561 CACAGGCTTCAGGGCTCTCTCTCACCGCAAGGGCAATTTCAAACTCAATGGCAAGTTCT 1620
QY 1669 CCGCAGCAGCTTAGAAGGCACTACCCCTAGCACTTTGGTCTCTCTGGAACAACTGGCCT 1728
Db 1621 CCGCAGCAGCTTAGAAGGCACTACCCCTAGCACTTTGGTCTCTCTGGAACAACTGGCCT 1680
QY 1729 CTTCCCATCTGAGCCCGCCAGCCGCTCAGGGGCTGTGAGTGAGGACAGCATCC 1788
Db 1681 CTTCCCATCTGAGCCCGCCAGCCGCTCAGGGGCTGTGAGTGAGGACAGCATCC 1740
QY 1789 TGTCTCTCGAGTCTTTTGACCAATTGGACTTGTGCTGAACTCTTCCGAAACCCCACTGA 1848
Db 1741 TGTCTCTCGAGTCTTTTGACCAATTGGACTTGTGCTGAACTCTTCCGAAACCCCACTGA 1800
QY 1849 GGGGCTGTGTCTGTGGAACAACTGAGGGGGTTGAGCAGCTCTCTCAGAAAGGTCTGA 1908
Db 1801 GGGGCTGTGTCTGTGGAACAACTGAGGGGGTTGAGCAGCTCTCTCAGAAAGGTCTGA 1860
QY 1909 AGCGATGGTGGCAGGAATCTTGGGGATAGCTGCTTTTCTCTGACAGACTGCCAAGAG 1968
Db 1861 AGCGATGGTGGCAGGAATCTTGGGGATAGCTGCTTTTCTCTGACAGACTGCCAAGAG 1920
QY 1969 TGAATGAGCTCAGACAAAGCCCTAGGAATCTGCTCAAAAGCTCAGCTGAGGAAGGAGA 2028
Db 1921 TGAATGAGCTCAGACAAAGCCCTAGGAATCTGCTCAAAAGCTCAGCTGAGGAAGGAGA 1980
QY 2029 TGGTGCCCTAGTAGGGGTAGGCTCTGAGAGGGTTTGCAGAGAAACC 2075
Db 1981 TGGTGCCCTAGTAGGGGTAGGCTCTGAGAGGGTTTGCAGAGAAACC 2027
```

RESULT 5

US-10-322-281-21

; Sequence 21, Application US/10322281

```
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 37278
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(37278)
; OTHER INFORMATION: n = A,T,C or G
; US-10-322-281-21
```

```
Query Match 66.8%; Score 1938.8; DB 17; Length 37278;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1940; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 955 CCGATGCCCTGTGGCCTGATCCGGTGGCTGTAAATGGTGAACCCACCCGTCGGGCCACAC 1014
Db 25129 CAGATGCCCTGTGGCCTGATCCGGTGGCTGTAAATGGTGAACCCACCCGTCGGGCCACAC 25188
QY 1015 TGGAGATGTAGCAGTCAATTGGTGGGTCAACTGGGGTTACACACCGGAGTGGGGAAC 1074
Db 25189 TGGAGATGTAGCAGTCAATTGGTGGGTCAACTGGGGTTACACACCGGAGTGGGGAAC 25248
QY 1075 AGGAAGCCCTGCGTGGAGGTGGGCACCTAGTGGTGAATTTGGCCGGGCTCCATGGCGG 1134
Db 25249 AGGAAGCCCTGCGTGGAGGTGGGCACCTAGTGGTGAATTTGGCCGGGCTCCATGGCGG 25308
QY 1135 ACTGGTTACGTGCTCTCTCGCCCTCTCTGGAGAAATGGAGCAAGGTGTGACGTTCT 1194
Db 25309 ACTGGTTACGTGCTCTCTCGCCCTCTCTGGAGAAATGGAGCAAGGTGTGACGTTCT 25368
QY 1195 TCAAGCAGACCTGCGCGGAGGTGGAGCACTGTACTGGGCTGGAGCGGCAATTTCTC 1254
Db 25369 TCAAGCAGACCTGCGCGGAGGTGGAGCACTGTACTGGGCTGGAGCGGCAATTTCTC 25428
QY 1255 TTAAGAAGTCCCGAAAGGAGATGACATGCTCAAAATCTGCAAGGTGACCCGGCTGAGG 1314
Db 25429 TTAAGAAGTCCCGAAAGGAGATGACATGCTCAAAATCTGCAAGGTGACCCGGCTGAGG 25488
QY 1315 ATACCTCTTCTCGCCCTGGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAA 1374
Db 25489 ATACCTCTTCTCGCCCTGGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATTTCTCAAGAAA 25548
QY 1375 AGTCTCTACCTGTCAGGGAGGTACAGGAGACCTCAGGAACCTCAGAACCGGTGCTG 1434
Db 25549 AGTCTCTACCTGTCAGGGAGGTACAGGAGACCTCAGGAACCTCAGAACCGGTGCTG 25608
QY 1435 ATACTCCAGGGCAGCTGCTCCCTGCTGTATCCCTGCTCCCAAGAAAGGCAATCTTAAAGA 1494
Db 25609 ATACTCCAGGGCAGCTGCTCCCTGCTGTATCCCTGCTCCCAAGAAAGGCAATCTTAAAGA 25668
QY 1495 AGTCTCGACAGCGTGAATCTGGTTACTTCTCTCCAGAGCCAGCAGTCTGGGGAAC 1554
Db 25669 AGTCTCGACAGCGTGAATCTGGTTACTTCTCTCCAGAGCCAGCAGTCTGGGGAAC 25728
QY 1555 TCTTAGACCGCAGTGTGTGTGAGTGGGACCCCGTGGAGCAGAAAGTCTTCCACAGG 1614
Db 25729 TCTTAGACCGCAGTGTGTGTGAGTGGGACCCCGTGGAGCAGAAAGTCTTCCACAGG 25788
QY 1615 CTTTCCGGCTCTCTCCCAAGGCAATTTCTCAAACTCAATGGAAGTCTTCCCGCA 1674
Db 25789 CTTTCCGGCTCTCTCCCAAGGCAATTTCTCAAACTCAATGGAAGTCTTCCCGCA 25848
QY 1675 CAGCCTTAGAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCCTCTCTCC 1734
```



Db 25849 CAGCTTAGAGGACATACCCCTAGACACCTTTGGCTCCCTGGACCAACTGGCCCTCCCTCCC 25908  
Qy 1735 ATCTGTGAGCCGCGCCAGCGCCCTCAGGGGCTGTGAGTGAGCAGCATCTCTGTCTT 1794  
Db 25909 ATCTGTGAGCCGCGCCAGCGCCCTCAGGGGCTGTGAGTGAGCAGCATCTCTGTCTT 25968  
Qy 1795 CCGAGTCTCTTGACCAATTTGGACTTGGCTGACCTTCCCGAAACCCCACTGAGGGGCT 1854  
Db 25969 CCGAGTCTCTTGACCAATTTGGACTTGGCTGACCTTCCCGAAACCCCACTGAGGGGCT 26028  
Qy 1855 GTGTGTCTGTGGACAACTGAGGGGGCTTGAGCAGCCTCCCTCAGAAGGCTCTGAAGCGAT 1914  
Db 26029 GTGTGTCTGTGGACAACTGAGGGGGCTTGAGCAGCCTCCCTCAGAAGGCTCTGAAGCGAT 26088  
Qy 1915 GTGTGGCAGGAATCTCTGGGGATAGCTGCTTTTCTCTGACAGACTGCCAAGAGTGACTG 1974  
Db 26089 GTGTGGCAGGAATCTCTGGGGATAGCTGCTTTTCTCTGACAGACTGCCAAGAGTGACTG 26148  
Qy 1975 CAGCTACAGACAGCCCTAGGAATCTGCTCAAGCTCAGCTGAGGAGGAGATGTGTC 2034  
Db 26149 CAGCTACAGACAGCCCTAGGAATCTGCTCAAGCTCAGCTGAGGAGGAGATGTGTC 26208  
Qy 2035 CTTAGTAGGGTAGGCTCTCAGAGGGTTTGCAGAGGAACCTCGGTGCGATTCCTCCAG 2094  
Db 26209 CTTAGTAGGGTAGGCTCTCAGAGGGTTTGCAGAGGAACCTCGGTGCGATTCCTCCAG 26268  
Qy 2095 TGAATAGAGTACATCAAGGGCTCTACGTCTGACGCTGAGTGAACCTGAAAGATGAGAGA 2154  
Db 26269 TGAATAGAGTACATCAAGGGCTCTACGTCTGACGCTGAGTGAACCTGAAAGATGAGAGA 26328  
Qy 2155 AATCGCATTCATGTGGAAGGAATGGGAACCTTGTCTGCCGAGTGTTATAGTGGGGTG 2214  
Db 26329 AATCGCATTCATGTGGAAGGAATGGGAACCTTGTCTGCCGAGTGTTATAGTGGGGTG 26388  
Qy 2215 CTTGAAGGTGCTACCTCTTTGTGCCATGAGTGACCCATGACATTTCCCAACCTTGTT 2274  
Db 26389 CTTGAAGGTGCTACCTCTTTGTGCCATGAGTGACCCATGACATTTCCCAACCTTGTT 26448  
Qy 2275 CTCTGGGTGACCTTACATAAGTTCTTGTTCATCAACACACAGGGTTAGAACCTCTGA 2334  
Db 26449 CTCTGGGTGACCTTACATAAGTTCTTGTTCATCAACACACAGGGTTAGAACCTCTGA 26508  
Qy 2335 CTTCTGGGAGGTAATGTGTAGTGACCTGCCATTTATTTAGAGAGGAACACAGCTCTGTTT 2394  
Db 26509 CTTCTGGGAGGTAATGTGTAGTGACCTGCCATTTATTTAGAGAGGAACACAGCTCTGTTT 26568  
Qy 2395 CCATCTCTGTGTGTCATCTCAAGACTGGGAAGACTCGGACCGCTGTTTGACTTCA 2454  
Db 26569 CCATCTCTGTGTGTCATCTCAAGACTGGGAAGACTCGGACCGCTGTTTGACTTCA 26628  
Qy 2455 TCTCAAGGGGACAGATGCCCCCTGGACCCCATCTTAGATCTCAGAGACTTGAACCTTGAA 2514  
Db 26629 TCTCAAGGGGACAGATGCCCCCTGGACCCCATCTTAGATCTCAGAGACTTGAACCTTGAA 26688  
Qy 2515 GCTGTTCTTAGTACCCAGATGTGATGATGCTGTTTCTCAGGCCAACGGACCTTAGA 2574  
Db 26689 GCTGTTCTTAGTACCCAGATGTGATGATGCTGTTTCTCAGGCCAACGGACCTTAGA 26748  
Qy 2575 ATGTGTGACTTATTTATTTTGTGTGATCTCACTCTCTTTTTTGTGTTTGTGTTT 2634  
Db 26749 ATGTGTGACTTATTTATTTTGTGTGATCTCACTCTCTTTTTTGTGTTTGTGTTT 26808  
Qy 2635 TGTGTTGTTTTGTTTTTAAGTGAATTTGCTGCTTTCAATAATGTAAGTCTGTGTTCTG 2694  
Db 26809 TGTGTTGTTTTGTTTTTAAGTGAATTTGCTGCTTTCAATAATGTAAGTCTGTGTTCTG 26868  
Qy 2695 GGGAACTCCACTGTCACCTGAAGTTTATGTACAGAGAAATTTGGCAATGATGTCCT 2754  
Db 26869 GGGAACTCCACTGTCACCTGAAGTTTATGTACAGAGAAATTTGGCAATGATGTCCT 26928  
Qy 2755 CTATTCAAGGGGGTGGGGGGTTTTTCAATGTATGTCTTGAGCACTGTCTGGATTGAG 2814

Db 26929 CTATTCAAGGGGGTGGGGGGTTTTTCAATGTATGTCTTGAGCACTGTCTGGATTGAG 25988  
Qy 2815 TCTCCAGTCCCTTTCACACCAAGGCTGGCCACCCCTCCCTCATCTTTCATCTGTGGCCAAA 2874  
Db 26989 TCTCCAGTCCCTTTCACACCAAGGCTGGCCACCCCTCCCTCATCTTTCATCTGTGGCCAAA 27048  
Qy 2875 AAAAAAAAAAAAAAAAAAAAAA 2896  
Db 27049 AAAAAAAAAAAAAAAAAAAAAA 27070

RESULT 6  
US-10-343-514-27  
; Sequence 27, Application US/10343514  
; Publication No. US20040132025A1  
; GENERAL INFORMATION:  
; APPLICANT: DRUCKER, Daniel J.  
; APPLICANT: ROSEN, Cheryl F.  
; APPLICANT: LEPREVRE, Diana L.  
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK  
; FILE REFERENCE: DPA-DRUC2/PCT  
; CURRENT APPLICATION NUMBER: US/10/343,514  
; CURRENT FILING DATE: 2003-01-31  
; PRIOR APPLICATION NUMBER: PCT/CA01/01109  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR APPLICATION NUMBER: US 60/222,650  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/274,613  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: CA 2,340,780  
; PRIOR FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 109  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 27  
; LENGTH: 2026  
; TYPE: DNA  
; ORGANISM: RAT  
US-10-343-514-27

Query Match 61.0%; Score 1771.4; DB 17; Length 2026;

Best Local Similarity 93.1%; Pred. No. 0;

Matches 1888; Conservative 0; Mismatches 136; Indels 5; Gaps 3;

Qy 49 GTGACCTCTGAGCCGCGGCTCAGCGCGGCTGCTACTGCTCCCGCCACCATCTCCACCTC 108  
Db 1 GTGACCTCTGAGCCTCGGCTCTCCGCGGCTGCTGCTCCCGCCACCATCTCCGCTC 60  
Qy 109 GCGGTCCCGCCACCATGAGTCTGGTGGCTTACTCCAGCGCCGCGGACCGGCTCCCTCGG 168  
Db 61 GCGGTCCCGCCACCATGAGTCTGGTGGCTTACTCCAGCGCGCGGGAACCTGGCTCCCTCGG 120  
Qy 169 CTTCCGCTCTGGCTCGGAGAGCGCGGCTGCGGACCGGCTCATCAAGTCCGCTA 228  
Db 121 CTTCCGCTCTGGCCACGAGAGCGCGGCTGCGGACCGGCTCATCAAGTCCGCTA 180  
Qy 229 AACCTCTGATGAAGAGCAGGCGGTGAAGCGCACCATCACAACACACCTGCGGCACC 288  
Db 181 AACCTCTGATGAAGAGCAGGCGGTGAAGCGCACCATCACAACACACCTGAGGACCC 240  
Qy 289 GCTACGAGTTCCTGGAGACGCTGGGCAAGGCGACCTTACGCGGAGGTGAAGAGGACGAG 348  
Db 241 GCTACGAGTTCCTGGAGACCCCTGGGCAAGGCGACCTTACGCGGAGGTGAAGAGGACGAG 300  
Qy 349 AGAGCTCGGGGCTCTGGTGGCCATCAAGTCCATCAGGAACACAAATCAAGATGAGC 408  
Db 301 AGAGCTCGGGGCGCTCTGGTGGCCATCAAGTCTATCAGGAAGGACAAATCAAGATGAGC 360  
Qy 409 AGGATCTGTCACATACGAGGAGATGAGATCATGTCTTCACTCAACACCCCCACA 468  
Db 361 AGGATCTGTTGACATAGGAGGAGATCGAGTCAATGTCTTCACTCAACACCCCCACA 420  
Qy 469 TCATTGCCCATCATGAGTGTGTGAATGAGCAGCAAGTGTGATTGTCTGAGGATGAG 528

Db 421 TCATTGCCATCCATGAAGTGTGTTGAGAAACAGCAAGAATTGTGATTGTCATGAGTACG 480  
Qy 529 CCAGCCGAGGCGATCTGTATGATTACATCAGTAGCGCGCCACAGCGCTCAGTGAGCGGACG 588  
Db 481 CCAGCCGAGGCGATCTGTACGATTACATCAGTAGCGCGCCACAGCGCTCAGTAGCGGACG 540  
Qy 589 CCAGGCAATTTCTCCGACAGATCGTGTCTGCGCTCGCACTACTGCGCAACAGGAGTCG 648  
Db 541 CCAGGCAATTTCTCCGACAGATCGTGTCTGCGCTCGCACTACTGCGCAACAGGAGTCG 600  
Qy 649 TTCAACCCAGATCTCAAGCTGGAAACATCTCTCTAGATGCCAATGGAAACATCAAGATTG 708  
Db 601 TTCAACCCAGATCTCAAGCTGGAAACATCTCTCTAGATGCCAATGGAAACATCAAGATTG 660  
Qy 709 CTGACTTTGGGCTCTCAACCTGTACCAACAGCAAGTTCTCTCAGACGTTCTGTGGGA 768  
Db 661 CTGACTTTGGGCTCTCAACCTGTATCAAAAGCGAAGTTCTCTCAGACGTTCTGTGGGA 720  
Qy 769 GCCCTCTCTACGCTCGCCTGAGATAGTCAACGGGAAAGCCCTATGTGGGCCAGAGGTGG 828  
Db 721 GCCCTCTCTACGCTCGCCTGAGATAGTCAACGGGAAAGCCCTATGTGGGCCAGAGGTGG 780  
Qy 829 ACAGCTGCTCTGCGGCGTTCTCTGTACATCTCTGTGATGCGACCATGCCCTTTGACG 888  
Db 781 ACAGCTGCTCTGCGGCGTTCTCTGTACATCTCTGTGATGCGACCATGCCCTTTGACG 840  
Qy 889 GGCAGGATCATAAACACTGTGTGAAGCAAACTCAGTAAACGGGCTTTACCGTGAGCGGCCCA 948  
Db 841 GGCAGGATCATAAACCTGTGTGAACAAATCAGTAGCGGGCTTTACCGTAGCGCGGTGCA 900  
Qy 949 AGCGCTCCGATGCTGTGGCTGTATCCGCTGATCCGCTGTTAAATGGTGAACCCACCGCTCGG 1008  
Db 901 AACCGTCTGATGCTGTGGCTGTATCCGCTGATCCGCTGTTAAATGGTGAATCCCATCCGTCGG 960  
Qy 1009 CCACACTGGAGATGTAGCAGTCACTGTGGTCACTGGGTACACCCGAGTCG 1068  
Db 961 CCACACTGGAGATGTAGCAGTCACTGTGGTCACTGGGTACACCCGAGTCG 1020  
Qy 1069 GGGAAACAGGAGCCCTCGCTGAGGCTGGGCAACCTAGTGTGACTTTGGCGGGCCCTCCA 1128  
Db 1021 GGGNACAGGAGCTCTCGAGAGGGTGGGCAACCTAGCGTACTCTGGCGGGCCCTCTA 1080  
Qy 1129 TGGCGGACTGTTACGTGCTCTCGCGCCCTCTCTGGAGAAATGGAGCCAAAGTGTGCA 1188  
Db 1081 TGGCGGACTGTTACGTGCTCTCGCGCCCTCTCTGGAGAAATGGAGCCAAAGTGTGTA 1140  
Qy 1189 GCTTCTTCAAGCAGCAGTGTGCGGAGGTGGAGCACTGTACCTGGGCTGGAGCGGCAAC 1248  
Db 1141 GCTTCTTCAAGCAGCAGTGTGCGGAGGTGGAGCACTGTACCTGGGCTGGAGCGGCAAC 1200  
Qy 1249 ATTCTCTTAAGAAAGTCCCGAAAGGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGG 1308  
Db 1201 ATTCTCTTAAGAAAGTCCCGAAAGGAGAAATGACATGGCTCAGACTCTGCAAGTGAACCCAG 1260  
Qy 1309 CTGAGGATACTCTTCTCGCCCTGGCAAGCAGCCTTAAGCTTCGAAAGGCATTCTCA 1368  
Db 1261 TTGAAGATACTCTCTCTCGCCCTGGCAAGCAGCCTCAAGCTTCGAAAGGTATCTCTCA 1320  
Qy 1369 AGAAAGTCTCTACTCTCTCAGGAGGTACAGGAGGACCTCAGGAATCAGACCGG 1428  
Db 1321 AGAAAGTCTCTCTCTCTCAGGAGGTACAGGAGGACCTCAGGAATCAGACCGG 1380  
Qy 1429 TGCTGTATCTCAAGGCGAGCTGTCTCTGCTGTATCTCTCTGCTCCCAAGGAAAGGCATCC 1488  
Db 1381 TGTCCAATACCCAGGCGAGCTGTCTCTGCTGTATCTCTCTGCTCCCAAGGAAAGGCATTC 1440  
Qy 1489 TTAAGAAAGTCTCGACAGCGTGAATCTGTGTTACTACTCTCTCAGAGGCCAGGAGTCG 1548  
Db 1441 TTAAGAAAGTCTCGACAGCGTGAATCTGTGTTACTACTCTCTCAGAGGCCAGGAGTCG 1500  
Qy 1549 GGGAACTCTTAGAGCCGAGTGTGTGTTGAGTGGGGACCCCGTGGAGCAGAAATCTC 1608  
Db 1501 GGGAACTCTTAGAGCCGAGTGTGTGTTGAGTGGGGACCCCGTGGAGCAGAAATCTC 1560

Qy 1609 CACAGGCTTCAGGCTCTCTCTCACCGCAAGGSCATTTCTCAAACTCAATGGCAAGTTCT 1668  
Db 1561 CACAAGCTTCAGG--CGCTCATCGCAAGGSCATTTCTCAAACTCAATGGCAAGTTCT 1617  
Qy 1669 CCGGACAGGCTTTAGAGGCACTACCCCTAGCACTTTTGGCTCCCTGGACCAACTGGCCT 1728  
Db 1618 CCGGACAGGCTTTAGAGGCACTGCCCCCTAGCACTTTTGGCTCCCTAGACCAACTGGCCT 1677  
Qy 1729 CTTCCATCTCTGAGCGCCGCGCCAGCGCCCTCAGGGCTGTGAGTGAGGACAGCATCC 1788  
Db 1678 CCGCTCATCTCTACAGCCCGGCGCAGCGCTCCCTCGGAGCTGTGAGTGAGGACAGCATCC 1737  
Qy 1789 TGTCTCTCCGAGTCTTTTGACCAATTTGCACTTTGCTCTGAACGCTTTCCGAAACCCCACTGA 1848  
Db 1738 TGTCTCTCCGAGTCTTTTGACCAATTTGCACTTTGCTCTGAACGCTTTCCGAAACCCCACTGA 1797  
Qy 1849 GGGGCTGTGTGTCTGTGGCACAACCTGAGGGGGCTTGAGCAGCTCCCTCAGAAAGGTCTGA 1908  
Db 1798 GAGCTGTGTGTCTGTGGCACAACCTGAGGAGGCTTGAGCAGCTCCCTCAGAAAGGCTAA 1857  
Qy 1909 AGCGATGTTGGCAGGAATCTTTGGGGATAGCTTTTCTGACAGACTGCCAAGAGG 1968  
Db 1858 AACGATGTTGGCAGGAATCTTTGGGGATAGCTTTTCTGACAGACTGCCAAGAGG 1917  
Qy 1969 TGACTGAGCTTACAGACAAAGCCCTAGGAATCTGCTCAAAAGCTCAGCTGAGGAAGGAGA 2028  
Db 1918 TGACAGAGCTTACAGACAAAGCCCTAGGAATCTGCTCGAAGCTCAGCTGAGGAGGAGG 1977  
Qy 2029 TGGTGCCCTAGT-ATGGGTAGGCTCTGAGA-GGGTTTGAGAGGAAC 2075  
Db 1978 CAGTGCCCGAGTGTAGGGTAGACTCTTAGAGGGGTTTGAGAGGAAC 2026

## RESULT 7

US-09-963-159-1  
; Sequence 1, Application US/09963159  
; Patent No. US20020077312A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Galvin, Katherine M.  
; TITLE OF INVENTION: 3700. A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR  
; FILE REFERENCES: 10147-50U1  
; CURRENT APPLICATION NUMBER: US/09/963,159  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 60/234,922  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 3353  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (3268)  
; NAME/KEY: unsure  
; LOCATION: (3270)  
; NAME/KEY: unsure  
; LOCATION: (3272)  
US-09-963-159-1

Query Match 47.5%; Score 1379; DB 9; Length 3353;  
Best Local Similarity 81.8%; Pred. No. 0;  
Matches 1654; Conservative 0; Mismatches 335; Indels 34; Gaps 4;  
Qy 92 CTACTGCTGCCCCGACCCACTCCACCTCGGGGTCCCGGACCACTGAGTGGTGGCTTAC 141  
Db 116 CTACTGATTTCCCTGCGGCCCTTGTCTACCTCTGCTCGCATGAGTGTGTTTCG 175  
Qy 142 TCAGGCGCGGAGCCAGGCTCCCTCGGCTTCGCGCTCGGAGAGGCGCGCGCGC 201  
Db 176 CGCGGCGCTCCGCGCCCACTCCCTCGGCGCGCAGAGCTA-----GCCGCGCGC 223

QY 202 TGGCGGCGGCTCATCAAGTCGCTTAACTCTGATGAAGACGAGCGGTGAAGCGC 261  
Db 224 TGGCGGAGGGCTGATCAAGTCGCTTAACTGAGCGGCTTAACTGAAGACGAGCGGTGAAGCGG 283  
QY 262 ACCATCAAAACACAACTGGGCGACCGCTACGAGTTCCTGGAGACGCTGGGCAAGGCA 321  
Db 284 ACCACCAAGCAACAACTGGGCGACCGCTACGAGTTCCTGGAGACCGCTGGGCAAGGCA 343  
QY 322 CTTAGGGAAGGTGAAGAGGCAAGAGAGCTGGGCGCTGCTGGTGGCCATCAAGTCCA 381  
Db 344 CTTAGGGAAGGTGAAGAGGCGGAGAGCTGGGCGCTGCTGGTGGCCATCAAGTCAA 403  
QY 382 TCAGGAAGACAAATCAAGATGAGCAGGAGTCTGTCACATACGAGGAGAGATTGGA 441  
Db 404 TCCGGAAGACAAATCAAGATGAGCAGGAGTCTGATGCAATACGAGGAGAGATTGGA 463  
QY 442 TCATGTCTTCACTCAACCAACCCACATCATTTGCCATTCATGAAGTGTGAGAACAGCA 501  
Db 464 TCATGTCTTCACTCAACCAACCCACATCATTTGCCATTCATGAAGTGTGAGAACAGCA 523  
QY 502 GCAAGATGTGATTCATGAGTATGCCAGCGAGGCGATCTGTATGATTAATCAAGTG 561  
Db 524 GCAAGATGTGATTCATGAGTATGCCAGCGGCGAGCTTTATGACTTACATCAGOG 583  
QY 562 AGCGGCGACGCTGAGTGAGCGGAGCGGACGAGCTTTCTCCGACAGATGTCTGCTGCC 621  
Db 584 AGCGGCGACGCTGAGTGAGCGGAGCGGAGCTTTCTCCGACAGATGTCTGCTGCC 643  
QY 622 TGCACTACTGCCACAGAGCGGATCGTTCAACGAGATCTCAAGCTGGAAACATCCTTC 681  
Db 644 TGCACTACTGCCACAGAGCTGTGTCACCGAGATCTCAAGCTGGAGAACATCCTCT 703  
QY 682 TAGATGCCATGGAACATCAAGATGCTGAGCTTTGGCTCTCAACCTGTACCAAAAG 741  
Db 704 TGGATGCCATGGAATATCAAGATGCTGAGCTTTGGCTCTCAACCTGTACCAAAAG 763  
QY 742 GCAAGTTCCTCCAGACGTTCTGTCGGAGCGCTCTACGCTCGCTGAGATAGTCAAG 801  
Db 764 GCAAGTTCCTCCAGACATCTGTCGGAGCGCTCTATGCTTCGCGAGATGTCAGT 823  
QY 802 GGAAGCCTATGTGGGCGGAGGTGACAGCTGCTCTGGGCGTTCTCTGTACATCC 861  
Db 824 GGAAGCCTATCACAGGCGGAGGTGACAGCTGCTCTGGGCGTTCTCTGTACATCC 883  
QY 862 TGTGATGACGACCATGCTTTTACGGGAGGATCATAAACACTGCTGAGCAATCA 921  
Db 884 TGTGATGACGACCATGCTTTTATGGGATGACCATTAAGATCTTAGTGAACAGATCA 943  
QY 922 GTAAAGGGGCTTACCGTGAGCGCGCCCAAGCGCTCGATGCTGTCGCTGATCCGCTGGC 981  
Db 944 GCAAGCGGCGCTACCGGAGCGCACCTAAACCTCTGATGCTGTCGCTGATCCGCTGGC 1003  
QY 982 TGTTAATGTTGAACCCCAACCGCTCGGCGCACTGTGAGGATGTAGCAGTCAITGTTGG 1041  
Db 1004 TGTGATGTTGAACCCCAACCGCGGCGCACCTGTGAGGATGTGGCCAGTCACTGTTGG 1063  
QY 1042 TCAACTGGGTTTACACACCGGAGTGGGGAACAGGAGCGCTGCTGAGGTTGGGAC 1101  
Db 1064 TCAACTGGGCTACCGGAGCGCACCTAAACCTCTGATGCTGTCGCTGATCCGCTGGC 1123  
QY 1102 CTAGTGTGACTTTGGCGGCGCTTCCATGCGGAGTGTGTTACGTCGCTCTCGCGCCCC 1161  
Db 1124 CTGCGAGTACTCTGCGCGGCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1183  
QY 1162 TCTGGAAGATGAGGCAAGGTGTGAGCTTTTCAAGCAGACGCTGGCGGAGGTGGA 1221  
Db 1184 TCTGGAAGATGAGGCAAGGTGTGAGCTTTTCAAGCAGACGATGCACTGTTGGGGA 1243  
QY 1222 GCACTGTACCTGGGCGGAGCAATCTCTTAAAGAGTCCCGAAGAGATGACA 1281  
Db 1244 GCACACCCCTGGCTGGAGCGCGAGCATTCGCTCAAGAAAGTCCCGAAGAGATGACA 1303

QY 1282 TGGCTCAAAATCTGCAAGGTGACCGCGCTGAGGATACCTTCTCGCCCTGGCAAGAGCA 1341  
Db 1304 TGGCCCAAGTCTCTCCACAGTGACACGCGCTGATGACATGCGCTTGGCAAGAGCA 1363  
QY 1342 GCCTTAAGCTTCCGAAGGCAATTCAGAGAAAGTCTCTTACTCTGCTCAGGGGAGGTAC 1401  
Db 1364 ACCTAAGCTGCCAAGGGGCAATTCAGAGAAAGGTGCTCAGCTCTGCAAGAGGGGTAC 1423  
QY 1402 AGAGGACCTCTAGGAACTCAGACCGGTGCTGATCTCAGGGGCAAGCTTCTGCTGCTG 1461  
Db 1424 AGAGGACCTCTCGAGCTCAGCCCAATCTCTGAGCGCCAGGCGAGCTGTGCTACT 1477  
QY 1462 TATCCCTGCTCCCAAGGAAAGGCAATCTTAAGAAAGTCTCAGACGCTGAATCTGTTACT 1521  
Db 1478 ---CCCTGCTCCCAAGGAAAGGCAATCTCAAGAAAGCGCCAGCGAGCTGTGCTACT 1534  
QY 1522 ACTCCTCTCCAGACCGCGAGTCTTGGGAACTCTTTAGACGCGAGTGAATCTGTTGTA 1581  
Db 1535 ACTCCTCTCCAGACCGCGAGTGAATCTTGGGAGTCTTTGACGCGAGGAGCTGTTGTA 1594  
QY 1582 GTGGGACCGCGTGGAGAGAAAGTCTCCACAGGCTTCAAGGCTCTCTCCACCGCAAG 1641  
Db 1595 GTGGGATCCCAAGGAGCAGAGGCTCTCGCAAGCTTCAAGGCTCTCTCTCATCGCAAG 1654  
QY 1642 GCATCTCAAACTCAATGGCAAGTCTTCCAGACAGCTTAGAGCTCGCGGCGCCCA 1701  
Db 1655 GCATCTCAAACTCAATGGCAAGTCTTCCAGACAGCTTAGAGCTCGCGGCGCCCA 1714  
QY 1702 CTTTGGCTCTCCGACCAACTGCGCTCTCTCCATCTGTCAGCGCGCGCGCGCGCT 1761  
Db 1715 CTTTGGCTCTCCGACCAACTGCGCTCTCTCCATCTGTCAGCGCGCGCGCGCGCT 1774  
QY 1762 CAGGGCTGTGAGTGAAGAGCAGATCTCTGCTCTCGAGTCTTTGACCAATTTGACTTGC 1821  
Db 1775 CAGGGCTGTGAGGAGGAGCAGATCTCTGCTCTGAGTCTTTGACCAAGTCTGCTTGC 1834  
QY 1822 CTGAAGCTCTTCCGAAACCCCACTGAGGGGCTGCTGCTGTCGAGCAACCTGAGGGGC 1881  
Db 1835 CTGAAGCTCTTCCAGAGCGCCCACTGCGGGCTGCTGCTGTCGAGCAACCTGAGGGGC 1894  
QY 1882 TTGAGCAGCTCTCTCAGAAAG-----GTCTGAAGCGATGGTGGCAGGAATCCT 1929  
Db 1895 TTGAGCAGCGCGCTCAGAGGGCGCTGGAAGCTGCTGAGGGCGCTGCGCGAGGATCCTT 1954  
QY 1930 TGGGGATAGTCTCTTTCTGACAGACTGCAAGAGTGAAGTGAAGCTGAGGAGTGAAG 1989  
Db 1955 TGGGGAGAGTCTCTTTCTGACAGACTGCGGAGGAGTGAAGGAGTGAAGGAGTGAAG 2014  
QY 1990 CTTAGGAATCTGCTCAAGCTCAGCTGAGGAGGAGATGGTGGCTTAGTATGGGGTAG 2049  
Db 2015 CACTGAGGCTGCTCAAGCTCAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2073  
QY 2050 GCTCTGAGGGTGTTCAGAGGAACTCTGGGTGGGATTCCTCC 2092  
Db 2074 GCTCTGAGTGAAGTGTTCAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2116

## RESULT 8

US-10-423-543-43  
; Sequence 43, Application US/10423543  
; Publication No. US20040058355A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Libermann, Robana K.  
; APPLICANT: Hunter, John J.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Rudolph-Owen, Laura A.  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Tsai, Fong-fing  
; APPLICANT: Galvin, Katherine M.  
; APPLICANT: Chun, Miyoung  
; APPLICANT: Williamson, Mark J.



```
Qy 1582 GTGGGGACCCCGTGGAGCAGAAGTCTCCACAGGCTTCAGGGCTCCTCTCCACCGCAAGG 1641
Db 1595 GTGGGGATCCCAAGGAGCAGAAGCCTCGCAAGCTTCAGGGCTCCTCTCCATCGCAAG 1654
Qy 1642 GCATTCTCAAACTCAATGGCAAGTCTCCCGCAAGCTTTAGAGGCACTACCCCTAGCA 1701
Db 1655 GCATCCCTCAAACTCAATGGCAAGTCTCCCGCAAGCTTTAGAGGCTCCGCGCCCAACA 1714
Qy 1702 CCTTTGGCTCCCTGGACCAACTGGCTCCTCCATCTCGAGCCGCGGCCAGCCGCTT 1761
Db 1715 CCTTGGCTCCCTGGATGAATCGCCCCACCTCGCCCCCTGGCCCCGGGCGAGCCCTT 1774
Qy 1762 CAGGGCTGTGAGTGAGGACAGACTCCTGTCTCCGAGTCTCTTGACCAATGGACTTGC 1821
Db 1775 CAGGGCTGTGAGTGAGGACAGACTCCTGTCTCTGAGTCTCTTGACCAAGTGAATTCG 1834
Qy 1822 CTGAACGCTCTCCCGAAACCCCACTGAGGGGCTGTGTCTGTGGACAACTGAGGGGGC 1881
Db 1835 CTGAACGGCTCCAGAGCCCCCACTGGCGGGCTGTGTCTGTGGACAACTCAAGGGGC 1894
Qy 1882 TTGAGCAGCTCCCTCAGAAG-----GTCTGAGCGATGGTGGCAGGAATCCT 1929
Db 1895 TTGAGGAGCCCCCTCAGAGGGCCCTGGAAGCTCGCTGAGCGCTGGCGGAGGATCCTT 1954
Qy 1930 TGGGGATAGTGTCTTTCTGTACAGACTGCCAAGAGGTGACTGCGAGCTTACAGCAAG 1989
Db 1955 TGGGGAGCAGCTGTCTTTCTGTACAGACTGCCAAGAGGTGACAGCGACCTACCGACAG 2014
Qy 1990 CCCTAGGAATCTGTCAAAGCTGAGTGAAGAGGAGATGGTGCCTTAGTATGGGGTAG 2049
Db 2015 CACTGAGGGTGTGTCAAAGCTCACTGAGTGAAGTGGCAATGGCCCGAG-CCCGGTGAG 2073
Qy 2050 GCTCTGAGAGGTTTGCAGAGGAACCTGGGTGCGATTCTCTCC 2092
Db 2074 GCTCTGAGATGACAGTGGTTGACCCCGAGGGAGATGCGCTTC 2116
```

## RESULT 9

US-10-322-281-25

```
; Sequence 25, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 3404
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-25
```

```
Query Match 47.4%; Score 1375.8; DB 17; Length 3404;
Best Local Similarity 81.7%; Pred. No. 0;
Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;
```

```
Qy 82 CTACTGTGCCGACCACTCCACCTCGCGGTCCCGCACCACATGGAGTTCGGTGGCTTAC 141
Db 87 CTACTGATTCCCTGCGGCTTGTCTCACTCTGCTCGCATGGAGTTCGCTGTTTCG 146
Qy 142 TCAGGCGCCGAGCAGGCTCCCTCGGCTCCGCTCGGCTCGGAGAGCGCCGCGCGC 201
Db 147 CGGGGCGCTCGGCGCCCACTCCCTCGGCGCAGAGCTA-----GCCGCGCGC 194
Qy 202 TGGCGGACGGGCTCATCAAGTCGCTAAACCTCTGATGAAGAAGCAGGCGGTGAAGCGC 261
Db 195 TGGCGGAGGGGCTGATCAAGTCGCGCCCAAGCCCTTAATGAAGAAGCAGGCGGTGAAGCGC 254
Qy 262 ACCATCAAAACAAACCTCGGCGACCGCTACGAGTTCTCTGAGACGCTGGGCAAGGGCA 321
```

```
Db 255 ACCACACACAGCACAACTTCGGGCAACGCTACAGTTCCTGGAGACCTTGGGCAAGGCA 314
Qy 322 CTACCGGAAAGGTGAAGAAGGACGAGAGAGCTCGGGGCGTCTGGTGGCCATCAAGTCCA 381
Db 315 CTACCGGAAAGGTGAAGAAGGCGCGGAGAGCTCGGGGCGCTGGTGGCCATCAAGTCAA 374
Qy 382 TCAGGAAAGCAAAATCAAGATGACGAGATCTGTGCACATACGAGGAGATTGAGA 441
Db 375 TCCGGAAGGCAAAATCAAGATGAGCAAGATCTGATGCACATACGAGGAGGATTGAGA 434
Qy 442 TCATGTCTTCACTCAACCAACCCCAATCATTTCCCATCCATGAAGTGTTTGAGAATAGCA 501
Db 435 TCATGTCTTCACTCAACCAACCCCTCATCATTTCCCATCCATGAAGTGTTTGAGAACAGCA 494
Qy 502 GCAAGATTTGATGTTCATGGAGTATGCGAGCGGAGCGATCTGTATGATTAATCATAGTG 561
Db 495 GCAAGATCTGTATCGTTCATGGAGTATGCGAGCGGCGGAGCTTTATGACTACATCAGCG 554
Qy 562 AGCGGCGACGGCTGAGTGGCGGAGCGCCAGGCAATTTCTTCGACAGATCGTCTCTGCC 621
Db 555 AGCGGCGAGCAGCTCAGTGGAGCGGAGCTAGGCAATTTCTTCGCGCAGATCGTCTCTGCC 614
Qy 622 TGCACTACTGCCCAACCGGATCGTTCAACGAGATCTCAAGCTTGGAAAAATCCTTTC 681
Db 615 TGCACTATTGCCCATCAGAACAGAGTTGTCCACCGAGATCTCAAGCTTGGAGAAATCCTCT 674
Qy 682 TAGATGCCAATGGAACATCAAGATTTGCTGAATTTGGCTCTTCCAACTGTGTACCAAAAG 741
Db 675 TGGATGCCAATGGAATATCAAGATTTGCTGAATTTGGCTCTTCCAACTGTGTACCAATCAAG 734
Qy 742 GCAAGTTCTCCAGAGCTTCTGTGGAGCGCTCTCTACGCTCGCTGAGATAGTCAACG 801
Db 735 GCAAGTTCTTCAGAGCATTTCTGTGGAGCGCTCTCTATGCTTCGCGAGAGATTGTCAATG 794
Qy 802 GGAAGCCCTATGTGGGCCAGAGGTTGGACAGCTGGTCTCTGGGCGGTTCCTCTGTACATCC 861
Db 795 GGAAGCCCTTACACAGGCCAGAGGTTGGACAGCTGGTCCCTGGGTGTCTCTCTACATCC 854
Qy 862 TGGTGTATGGCACCATGCCCTTTTACGGGCGAGGATCATAAACATCTGGTGAAGCAATCA 921
Db 855 TGGTGTATGGCACCATGCCCTTTTATGGGCGATGACCATTAAGATCTCTAGTGAACAGATCA 914
Qy 922 GTAACGGGCTTACCGTGAGCGCCCAAGCGCTCGATGCTGTGGCTGTATCCGCTGGC 981
Db 915 GCAACGGGCTTACCGGGAGCCACCTTAACCTCTGATGCTGTGGCTGTATCCGCTGGC 974
Qy 982 TGTAAATGGTGAACCCCAACCCCTCGGGCCACACTGTGAGGATGTAGCCAGTCAATTTGG 1041
Db 975 TGTGATGTGAACCCCAACCCCGCGCGCCACCTCGGAGGATGTGGCCAGTCACTGTGG 1034
Qy 1042 TCAACTGGGGTTACACACCGGAGTCCGGGAAACAGGAACCCCTGTGAGGTTGGGCAAC 1101
Db 1035 TCAACTGGGGCTACGCGACCCCGAGGAGCAGGAGGCTCCGATGAGGTTGGGCAAC 1094
Qy 1102 CTAGTGTGCTTTCGCGGGCTCCATGCGCGAGCTGGTTCAGTTCGCTCTCTCGCGCCCC 1161
Db 1095 CTGGCAGTGAATCTGCGCGGCTCCATGCTGTGATCTGGCTCGGCTTCTCTCCGCCCC 1154
Qy 1162 TCTGAGGAATGGAGCAAGGTGTGAGCTTTCTTCAAGCAGCAGCTGCGGGAGGTGGAA 1221
Db 1155 TCTGAGGAATGGGCGCAAGGTGTGAGCTTTCTTCAAGCAGCAGTGCACCTGGTGGGGAA 1214
Qy 1222 GCATGTACTGGGCTGGAGCGGCAACATTTCTTTAAGAGTCCCGAAGGAGAAATGACA 1281
Db 1215 GCACCAACCCCTGGCTGGAGCGCGAGCATTCGCTCAAGAAGTCCCGCAAGGAGAAATGACA 1274
Qy 1282 TGGCTCAAAATCTGCAAGGTGACCGGGCTGAGGATACCTCTTCTCGCCCTGGCAAGAGCA 1341
Db 1275 TGGCCCAAGTCTCTCCAGTGAACGGCTGATGACACTGCCCCATCGCCCTGGCAAGAGCA 1334
Qy 1342 GCCTTAAGCTTCGAAAGGCATTTCTCAAGAAAAGTCTCTTACCTCTGTGAGGGAGGTAC 1401
```

Db 1335 ACCTCAAGCTGCCAAGGGCAATCTCAAGAAGAGGTGTGAGCCCTCTGCAGAGGGGTAC 1394  
Qy 1402 AGGAGGACCTCTCAGAACTCAGACCGGTGCTGTATCTCAGAGGACGCTGTCTCTCTG 1461  
Db 1395 AGGAGGACCTCTCAGAGCTCAGCCAAATCTCTGAGCCCAAGGACGCTGCC----- 1448  
Qy 1462 TATCCCTCTCCCAAGGAAAGGCACTCTTAAGAACTCTCAGACGCTGAATCTGGTTACT 1521  
Db 1449 ---CCCTGCTCCCAAGAGGGCAATCTCAAGAAGCCCGACAGCGGAGTCTGGTACT 1505  
Qy 1522 ACTCTCTCAGAGCCAGAGAGTCTGGGAACTCTTAGAGCCAGAGTGTGTGTGA 1581  
Db 1506 ACTCTCTCAGAGCCAGTGAATCTGGGAGCTCTTGGAGCGAGGAGCGTGTGTGA 1565  
Qy 1582 GTGGGGACCCGCTGAGCAGAGGTCTCCACAGGCTTCAGGCTCTCCTCTCCACCGAAG 1641  
Db 1566 GTGGGATCCCAAGAGCAGAGCTCTCGAAGCTTCAGGCTCTCCTCATCGCAAG 1625  
Qy 1642 GCATTCTCAAACTCAATGGCAAGTCTCCGCAAGCTTTAGAGGCACTACCCCTAGCA 1701  
Db 1626 GCATCCTCAAACTCAATGGCAAGTCTCCAGACAGCTTTGGAGCTCGCGGCCCAACCA 1685  
Qy 1702 CTTTGGCTCTCGGACCACTGCGCTCTCTCCATCTCGAGCCCGGCGCAGCGCCCT 1761  
Db 1686 CTTTGGCTCTCGGATGAATCTCGCCCACTCTCGCCCTCTGGCCCGGCGCAGCGCCCT 1745  
Qy 1762 CAGGGCTGTGAGTGAGGACAGCTCTGTCTCTCGAGTCTCTTGACCAATGGACTTC 1821  
Db 1746 CAGGGCTGTGAGTGAGGACAGCTCTGTCTCTCGAGTCTCTTGACCAATGGACTTC 1805  
Qy 1822 CTGAACGCTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGGACAACTCAGGGGGC 1881  
Db 1806 CTGAACGCTCTCAGAGCCCCCACTGCGGGCTGTGTCTGTGGACAACTCAGGGGGC 1865  
Qy 1882 TTGAGGACCTCTCTCAGAAG-----GTCTGAAGCGATGTGGCAGGAATCCT 1929  
Db 1866 TTGAGGACCTCTCTCAGAAGGCTCTGGAAGCTGCTGAGGCGCTGGCGGAGGATCCTT 1925  
Qy 1930 TGGGGGATAGCTGTCTCTCTGACAGACTCCAGAGGCTGACCTCAGAGCAAG 1989  
Db 1926 TGGGGGACAGCTGTCTCTCTGACAGACTCCAGAGGCTGACAGGCTTACCGACAG 1985  
Qy 1990 CCCTAGGAATCTGCTCAAAAGCTCAGCTCAGGAAGGAGATGGTGCCTTAGTATGGGTAG 2049  
Db 1986 CACTGAGGGTCTGCTCAAAAGCTCAGCTGAGTGGAGTAGGCAATGCCCCAG-CCCGGTAG 2044  
Qy 2050 GCTCTGAGAGGGTTTGACAGAGAAACCTCTGGGTCTGGATCTCTCC 2092  
Db 2045 GCTCTCAGATGAGCTGTGTTGACCCCGGAGGAGATGCTCTTC 2087

## RESULT 10

US-10-370-715B-639  
; Sequence 639, Application US/10370715B  
; Publication No. US20040258678A1

## GENERAL INFORMATION:

; Patin Docket Preview  
; APPLICANT: BODARY, SARAH C.  
; APPLICANT: CLARK, HILARY  
; APPLICANT: BRISDELL, HUNTE  
; APPLICANT: JACKMAN, JANET  
; APPLICANT: SCHOENFELD, JILL R.  
; APPLICANT: WILLIAMS, P. MICKEY  
; APPLICANT: WOOD, WILLIAM I.  
; APPLICANT: WU, THOMAS D.  
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune  
; TITLE OF INVENTION: Related Diseases  
; FILE REFERENCE: P1948R1-US  
; CURRENT APPLICATION NUMBER: US/10/370,715B  
; CURRENT FILING DATE: 2003-02-21  
; NUMBER OF SEQ ID NOS: 742  
; SEQ ID NO 639  
; LENGTH: 3443

; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-370-715B-639

Query Match 47.4%; Score 1374.2; DB 18; Length 3443;

Best Local Similarity 81.8%; Pred. No. 0;  
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

Qy 82 CTACTGCTGCCGACCCACCTCCACCTCTCGGGTCCCGCACCACTGGAGTGGTGGCTTAC 141  
Db 87 CTACTGATTCCTCTCGGCTTGTCTCACTCTGCTCGCCATGGAGTGGTGGTTCG 146  
Qy 142 TCAGGCGCCGAGCCAGGCTCCTCTCGGCTCTCGGCTCTCGGAGAGCGCCGCGCGC 201  
Db 147 CGCGGCTCTCGGCGCCACCTCTCTCTCGGCGCAGAGCTA-----GCCCGCGCGC 194  
Qy 202 TGGCGAGCGGCTCATCAAGTGGCTTAACCTCTGATGAAGAGCAGGCGGTGAAGCGGC 261  
Db 195 TGGCGAAGGGCTGATCAAGTGGCCCAAGCCCTTAATGAAGAACGAGCGGTGAAGCGC 254  
Qy 262 ACCATCACAAACACAACTCGGCGACCGCTTACGAGTTCCTTGGAGAGCGCTGGGCAAGGCA 321  
Db 255 ACCACACAAAGACAACTCGGCGACCGCTACGAGTTCCTTGGAGAGCCCTGGGCAAGGCA 314  
Qy 322 CTTACGGGAGGTGAAGAAAGGCAACGAGAGCTCGGGGCGTCTGGTGGCCATCAAGTCCA 381  
Db 315 CTTACGGGAGGTGAAGAAAGGCGCGGAGCTCGGGGCGCTGGTGGCCATCAAGTCAA 374  
Qy 382 TCAGGAAGCAAAATCAAAAGATGAGCAGGATCTGTGCACATACCGAGGAGATTGAGA 441  
Db 375 TCCGGAAGGCAAAATCAAAAGATGAGCAAGATCTGATGACATACCGAGGAGATTGAGA 434  
Qy 442 TCATGCTTCACTCAACCCCACTCATTTGCCATCCATGAAGTGTTTGAGAATAGCA 501  
Db 435 TCATGCTCATCTCAACCCCACTCATTTGCCATCCATGAAGTGTTTGAGAACAGCA 494  
Qy 502 GCAAGATTGTGATTTGTATGGAGTATGCCAGCGAGCGGATCTGTATGATTAATCATGTG 561  
Db 495 GCAAGATCGTGTATCGTATGGAGTATGCCAGCGCGGCGGACCTTTATGACTACATCAGCG 554  
Qy 562 AGGGGCAAGCTGTGAGTGAGGAGCGCCAGGCAATTTCTTCCGACAGATCTGTCTGCC 621  
Db 555 AGGGGCAAGCTGTGAGTGAGGAGCGGAGCTAGGCAATTTCTTCCGCGAGATCTGTCTGCC 614  
Qy 622 TGCACTATCTCCACAGAGCGGATCGTTCCAGCGAGATCTCAAGCTGGAAGAACATCTCTC 681  
Db 615 TGCACTATTTGCCATCAGACAGAGTGTTCACCGAGATCTCAAGCTGGAGAACATCTCTC 674  
Qy 682 TAGATGCCAATGGAAACATCAAGATTTGTGACTTTGGCTCTCCAACTGTACCAAAAG 741  
Db 675 TGGATGCCAATGGAAATATCAAGATTTGTGACTTTGGCTCTCCAACTGTACCAATCAAG 734  
Qy 742 GCAAGTTCTCAGAGCTTCTGTGGAGCGCTCTTACGCTCGCTGAGATAGTCAAG 801  
Db 735 GCAAGTTCTCAGAGCTTCTGTGGAGCGCCCTCTATGCTCGCCAGAGATTTGTCAATG 794  
Qy 802 GGAAGCCCTATGTGGGCGCAGAGTGTGACAGTGTCTCTGGGCTTCTCTGTATCATCC 861  
Db 795 GGAAGCCCTATACAGGCGCCAGAGTGTGACAGTGTGTTCCCTGGGTGTCTCTCTCATCC 854  
Qy 862 TGGTGCATGCGACCATGCCCTTTGACGGGAGGATCATAAACATCTGGTGAAGCAAAATCA 921  
Db 855 TGGTGCATGCGACCATGCCCTTTGATGGGCAATGACCATTAAGATCTTAGTGAACAGATCA 914  
Qy 922 GTAAAGGGCTTACGTTAGCGCCCGCCAGCGCTCGATGCTGTGGCTGATCCGTTGGC 981  
Db 915 GCAAGCGGGCTTACCGGGAGCCACCTAAACCCCTCTGATGCTGTGGCTGATCCGTTGGC 974  
Qy 982 TGTAAATGTGTAACCCCGCTCGGCGCACACTGGAGGATGTAGCCAGTCAATTTGGTGGG 1041  
Db 975 TGTGATGTGTAACCCCGCTCGGCGCACACTGGAGGATGTGGCGGATGTGGCTGGG 1034  
Qy 1042 TCAACTGGGGTTACACCAAGCGGAGTCCGGGAAACAGGAAGCCCTGCTGTGAGGTGGCACC 1101



```
Db 1035 TCAACTGGGCTACGCCACCCGAGTGGGAGACGAGGCTCCGATAGAGGTGGGACCC 1094
Qy 1102 CTAGTGGTGAATTTGGCCGGGCTCCATGGCGGACTGTTTACGTGCTCTCTCGCGCCCC 1161
Db 1095 CTGGCAGTGAATCTGCGCGGCTCCATGGCTGACTGGCTCCGGGTTCTCTCCGCCCC 1154
Qy 1162 TCCTGGGAATGGAGCCAGGTGTGACGCTTCTTCAAGCAGACAGTGGCCGGAGGTGGAA 1221
Db 1155 TCCTGGGAATGGGGCCAAAGTGTGACGCTTCTTCAAGCAGACATGCACCTGTGGGGAA 1214
Qy 1222 GCACCTGACCTGGGCTGGAGCGGCAATCTCTTAAGAACTCCCGAAAGAGATGACA 1281
Db 1215 GCACACCCCTGGCTGGAGCGCCAGCATCTGCTCAAGAACTCCCGAAGAGATGACA 1274
Qy 1282 TGGCTCAAAATCTGCAAGGTGACCGGCTGAGGATACCTCTTCTGCGCTGGCAAGAGCA 1341
Db 1275 TGGCCCACTCTCCACAGTGACACGGCTGATGACACTGCCCATCGCCCTGGCAAGAGCA 1334
Qy 1342 GCCTTAGCTTCGGAAGGCATCTCAAGAAAGTCTCTACCTGTCAGGAGGTAC 1401
Db 1335 ACCTCAAGCTCCAAAGGCAATCTCAAGAAAGGTGTGACGCTCTGCAAGAGGATAC 1394
Qy 1402 AGGAGGACCTTCAGGAACTCAGACCGGTGCTGATCTCCAGGCGAGCCTGTGCTGCTG 1461
Db 1395 AGGAGGACCTTCAGGAGCTCAGCCCAATCTCTGCGAGCCCAAGGAGGCTGCC- 1448
Qy 1462 TATCCTGCTCCAAAGAAAGCATCTCTTAAGAAAGTCTGCAAGCGTGAATCTGGTACT 1521
Db 1449 ---CGCTGCTCCCAAGAAAGGCATCTCAAGAAAGCCCGACAGCGGAGTCTGGCTACT 1505
Qy 1522 ACTCCTCTCCAGAGCCAGCGAGCTGGGGAATCTTTAGAGCCAGTGAATGTTTCTGA 1581
Db 1506 ACTCCTCTCCAGAGCCAGCGAGTCTGGGAGCTCTTGGAGCGAGGAGCTGTTTCTGA 1565
Qy 1582 GTGGGGACCCGCTGGAGCAGAAGTCTCCAGAGGCTTCAGGCTCTCTCTCCAGCGCAAG 1641
Db 1566 GTGGGGATCCCAAGAGCAGAAGCCTCGCAAGCTTCAGGCTCTCTCTCCAGCGCAAG 1625
Qy 1642 GCATCTCAAACTCAATGGGAAAGTCTTCCGCAAGAGCTTTAGAGGCACTACCCCTAGCA 1701
Db 1626 GCATCTCAAACTCAATGGGAAAGTCTTCCAGAGAGCTTTGGAGCTCGCGGCCCCACCA 1685
Qy 1702 CTTTGGCTCCCTGAGCAACACTGCGCTCTCTCCATCTGAGCCCGGCGGCGGCGCCCT 1761
Db 1686 CTTTGGCTCCCTGAGTGAATCTGCGCAACTCTGCGCCCTGCGCGGCGGCGGCGGAGCCCT 1745
Qy 1762 CAGGGCTGTGAGTGAGGACAGCATCTCTGCTCTCCGAGTCTCTTGAACCAATTTGACTTGC 1821
Db 1746 CAGGGCTGTGAGTGAGGACAGCATCTCTGCTCTGAGTCTCTTGAACCAATTTGACTTGC 1805
Qy 1822 CTGAACTCTTCCGAAACCCACTGAGGGGCTGTGTGTGTGGAACACTGAGGGGG 1881
Db 1806 CTGAACTCTTCCGAAACCCCACTGAGGGGCTGTGTGTGTGGAACCACTCAGGGGG 1865
Qy 1882 TTGAGCAGCTCTCTCAGAAG- 1929
Db 1866 TTGAGGAGCCCCCTCAGAGGGCCCTGGAAGTCCCTGAGGCGCTGGCGGAGGATCCCT 1925
Qy 1930 TGGGGGATAGTGTCTTCTCTGACAGACTGCCAAGAGGTGACTGACGCTTACAGACAAG 1989
Db 1926 TGGGGGACAGTGTCTTCTCTGACAGACTGCCAAGAGGTGACAGCGACCTACCGACAG 1985
Qy 1990 CCTTAGGAATCTGCTCAAGCTCAGCTGAGGAGGAGATGGTCCCTAGTATGGGGTAG 2049
Db 1986 CACTGAGGGTCTGCTCAAGAGCTCACTGAGTGGAGTAGGATTTGCCCCAG-CCCGGTGAG 2044
Qy 2050 GCTCTGAGAGGGTTTGACAGAGGAACCTGGGTGCGATTTCCTCC 2092
Db 2045 GCTCTGAGTGCAGCTGGTTGACACCCCGAGGGAGATGCGCTTC 2087
```

```
US-10-618-941-11
; Sequence 11, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 3463
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-618-941-11
```

```
Query Match 47.4%; Score 1374.2; DB 18; Length 3463;
Best Local Similarity 81.6%; Pred.No. 0;
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;
```

```
Qy 82 CTACTGCTGCCGACCCACTCCACCTCGCGGTCCCGCACCATGAGTGGTGGCTTAC 141
Db 148 CTATGATTCCCTGCGGCTTTGCTTCACCTCTGCTCGCCATGGAGTGGCTGTTTCG 207
Qy 142 TCCAGCGCCGAGCCAGGCTCCCTCGGCTCGGCTCGGAGAGCGCCGCGCCG 201
Db 208 CGGGGCTCGGCGCCACTCCCTCGGCGCAGAGCTA-----GCCGCGCGC 255
Qy 202 TGGCGAGCGGCTCATCAAGTCCCTTAAACCTCTGATGAAGAACGAGGCGGTGAAGCGGC 261
Db 256 TGGCGGAAGGGCTGATCAAGTCCGCAAGCCCTTAATGAAGAACGAGGCGGTGAAGCGGC 315
Qy 262 ACCATCACAAACACACCTCGGCGACCGCTACGAGTCTCTGAGAGCGCTGGGCAAGGGCA 321
Db 316 ACCACACAGCAACACCTCGGCGACCGCTACGAGTCTCTGAGAGCGCTGGGCAAGGGCA 375
Qy 322 CTTACGGGAAGGTGAAGAGGACGAGAGAGCTCGGGGCTCTGTTGGCCATCAAGTCCA 381
Db 376 CTTACGGGAAGGTGAAGAGGCGGGGAGAGCTCGGGGCGCTGTTGGCCATCAAGTCAA 435
Qy 382 TCAGGAAGACAAAATCAAAGATGAGCAGGATCTGCTGCACATACGAGGAGATTGAGA 441
Db 436 TCCGGAAGGACAAAATCAAAGATGAGCAAGATCTGATGCACATACGAGGAGATTGAGA 495
Qy 442 TCATGCTTTCACTCACACACCCCACTATTCGCCATCCATGAGTGTGTTGAGAAATAGCA 501
Db 496 TCATGCTATCACTCAACACCCCTCATTCATTTGCCATCCATGAAGTGTGTTGAGAAATAGCA 555
Qy 502 GCAAGATTGTGATGTGATGAGGAGTATGCGAGCGGAGGCTCTGTATGATTACATCAGTG 561
Db 556 GCAAGATTGTGATGTGATGAGGATATGCGAGCGGGGCGACCTTTATGACTACATCAGCG 615
Qy 562 AGCGGCCACGGCTGAGTGAAGCGGAGCGCCAGGCAATTTCTTCGACAGATGCTGTGTCTGCC 621
Db 616 AGCGGCAGCAGCTCAGTGAAGCGGAGCTAGGCAATTTCTTCGCGCAGATGCTGTCTGCCG 675
Qy 622 TGCACCTACTGCCACCAAGAGCGGATGTTTCCAGGATCTCAAGCTGGAACATCCTTTC 681
Db 676 TGACCTATTGGCATCAGAACAGAGTGTCTCCACCGAGATCTCAAGCTGGGAGAACATCCTCT 735
Qy 682 TAGATGCCAATGGAACATCAAGATTGCTGACTTTTGGCCCTCTCCAACTGTGTTGAGAAAG 741
Db 736 TGGATGCCAATGGGAATATCAAGATTGCTGACTTCGGCCCTCTCCAACTCTTACCATCAAG 795
Qy 742 GCAAGTTCCTCAGAGCTTCTGTGGAGGCCCTCTCTACGCCCTCGCTGAGATGATCAAGCG 801
Db 796 GCAAGTTCCTGACAGATTTCTGTGGAGGCCCTCTCTATGCTCGCCCTCGCAGAGATTGTCAATG 855
```



```
Qy 802 GGAAGCCCTATGTGGGCCAGAGGTGGAAGAGCTGGTCTCTGGGGGTTCTCTCTGTACATCC 861
Db |||||
Qy 856 GGAAGCCCTACACAGGCCACAGAGTGGACAGCTGGTCCCTGGGTGTTCTCTCTACATCC 915
Db |||||
Qy 862 TGGTGCATGGACCATGCCCTTTGACGGGACAGGATCATTAACACACTGGTGAAGCAATCA 921
Db |||||
Qy 916 TGGTGCATGGACCATGCCCTTTGATGGGCATGACCAATTAAGATCCTAGTGAACACAGATCA 975
Db |||||
Qy 922 GTAAAGGGGGTTACCGTGAAGCCGCCCAAGCCGTCCTGATGCTGTGGCTGATCCGGTGGC 981
Db |||||
Qy 976 GCACGGGGCTACCGGGAGCCACTAAACCTCTGATGCTGTGGCTGATCCGGTGGC 1035
Db |||||
Qy 982 TGTTAATGTGAACCCCAACCGTCCGGCCACACTGAGAGGATGTACAGCAGTCAATTTGGTGG 1041
Db |||||
Qy 1036 TGTGTAGTGTGAACCCCAACCGCCGGGCCACCTGGAGGATGTGCCAGTCACTGTGTGG 1095
Db |||||
Qy 1042 TCAACTGGGGTTACACACCGGAGTCGGGGAACAGGAAGCCCTGCTGAGGTGGGCACC 1101
Db |||||
Qy 1096 TCAACTGGGGCTACGCCACCGAGTGGGAGACAGGAGGCTCCGCATGAGGGTGGGCACC 1155
Db |||||
Qy 1102 CTAGTGTGACTTTGGCGGGCCCTCCATGGCGGACTGGTTACGTCGCTCTCGCGCCCCC 1161
Db |||||
Qy 1156 CTGCGAGTGACTTGGCGGCGCTCCATGGCTGACTGGCTCCGGCGTTCTCCCGCCCCC 1215
Db |||||
Qy 1162 TCCTGAGAAATGGAGCCAAAGTGTSCAGCTTTCTTCAAGCAGCAGCTGCGCGGAGGTGGAA 1221
Db |||||
Qy 1216 TCCTGGAGAATGGGGCCAAAGTGTGCAGCTTCTTCAAGCAGCATGCACCTGGTGGGGAA 1275
Db |||||
Qy 1222 GCATGTACTGGCTGGAGGGGGAATCTCTTTAAGAGTCCCGAAGAGAGATGACA 1281
Db |||||
Qy 1276 GCACCACTTGGCTGGAGCGCCAGCATTCGCTCAAGAAGTCCCGAAGAGAGATGACA 1335
Db |||||
Qy 1282 TGGCTCAAAATCTCAAGGTGACCCGCTGAGGATACCTCTTCTCGCCCTGGCAAGAGCA 1341
Db |||||
Qy 1336 TGGCCAGTCTCTCCAGTGAACAGGCTGATGACCTGCCATCGCCCTTGGCAAGAGCA 1395
Db |||||
Qy 1342 GCCTTAAGCTTCCGAAGGCATCTCAAGAAAAAGTCTCTTACCTCTGTCAGGGGAGGTAC 1401
Db |||||
Qy 1396 ACCTCAAGCTGCCAAAGGGCATCTCAAGAAGAGGTGTCAAGCTCTGCAAGAGGGGTAC 1455
Db |||||
Qy 1402 AGGAGGACCTCAGGAATCAGACCGGTGCTGATACTCCAGGGCAGCCTGTCCCTGCTG 1461
Db |||||
Qy 1456 AGGAGGACCTTCCGGAGCTCAGCCCAATCCCTGCGAGCCAGGCGAGGCTGCC- 1509
Db |||||
Qy 1462 TATCCCTGCTCCCAAGGAAGGCATCTTAAGAGTCTCCACAGCGTGAATCTGGTTACT 1521
Db |||||
Qy 1510 ---CCCTGCTCCCAAGAGGCAATCTCAAGAAGCCCGACAGCGGAGTCTGGTACT 1566
Db |||||
Qy 1522 ACTCCTCTCCAGAGCCAGCGAGTCTGGGGAACTCTTTAGAGCCAGTGTGTTGTGA 1581
Db |||||
Qy 1567 ACTCCTCTCCGAGCCAGTGAATCTGGGAGCTCTTGGAGCGAGGCGACGTGTTGTGA 1626
Db |||||
Qy 1582 GTGGGACCCCGTGGAGAGAGTCTCCAGAGCTTCAGGGCTCCTCTCCACCGCAAGG 1641
Db |||||
Qy 1627 GTGGGATCCCAAGAGCAGAGCTCCCGAAGCTTCAGGGCTGCTCTCCATCGCAAG 1686
Db |||||
Qy 1642 GCATTTCAACTCAATGGCAAGTTCTCCGCAAGCTTTAGAGGCACTACCCCTAGCA 1701
Db |||||
Qy 1687 GCATCTCAAACTCAATGGCAAGTTCTCCAGAGCTTTGAGCTCGCGGCCCCACCA 1746
Db |||||
Qy 1702 CTTTGGCTCCCTGGACCAACTGGGCTCTCCATCCATCTGAGCGCCGCCAGCGCCCT 1761
Db |||||
Qy 1747 CTTTGGCTCCCTGGATGACTCGCCCAACTCGCCCTTGGCCCGGCGCCAGCGACCT 1806
Db |||||
Qy 1762 CAGGGCTGTGAGTGAAGACAGATCTCTCTCCAGTCTTTGACCAATGGAATGTC 1821
Db |||||
Qy 1807 CAGGGCTGTGAGCGAGGACAGATCTCTCTCTGAGTCTTTGACCACTGAGCTTGC 1866
Db |||||
Qy 1822 CTGAACGTCTTCCGAACCCCACTGAGGGGCTGTGTCTGTGGACAACTGAGGGGGC 1881
Db |||||
Qy 1867 CTGAACGGCTCCCAAGAGCCCCCACTGCGGGGCTGTGTCTGTGGACAACTCAGGGGC 1926
Db |||||
Qy 1882 TTGAGCAGCCTCCCTCAGAAG-----GTCTGAAGCGATGTTGGCAGGAATCCT 1929
Db |||||
```

```
Db |||||
Qy 1927 TTGAGGAGCCCTCTCAGAGGGCCCTGGAAGCTGCCTGAGCGCTGCGCGCAGGATCCTT 1986
Db |||||
Qy 1930 TGGGGATAGCTGCTTTCTCTACAGACTGCCAAGAGGTGACTGACGCTTACAGACAG 1989
Db |||||
Qy 1987 TGGGGGACAGCTGCTTTCTCTACAGACTGCCAAGAGGTGACAGCAGCTTACCGCAGG 2046
Db |||||
Qy 1990 CCCTAGGAATCTGCTCAAACTCAGCTGAGGAAGGGAGATGGTGCCTAGTATGGGGTAG 2049
Db |||||
Qy 2047 CACTGAGGCTGCTCTCAAGCTCACCTGAGTGGAGTAGGATTTGCCCCAG-CCCGGTGAG 2105
Db |||||
Qy 2050 GCTCTGAGAGGGTTTGAGAGGAACCTCTGGGTGGATTCCTCC 2092
Db |||||
Qy 2106 GCTCTCAGATGCACTGGTTGACCCCGAGGGAGATGCTCTT 2148
Db |||||

RESULT 12
US-10-302-172-215
; Sequence 215, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids an
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 803.1CNP
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pt FL_genes Version 2.0
; SEQ ID NO 215
; LENGTH: 2501
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(1989)
US-10-302-172-215

Query Match 47.3%; Score 1372.2; DB 16; Length 2501;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1649; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

Qy 82 CTAAGTCTGCGGACCCACCTCCACCTCGCGGTCCCGCACCATTGGAGTCGGTGGCCTTAC 141
Db |||||
Qy 65 CTATTGATTCCTCGCGCCCTTGTCTCCTCTGCTCGCCTGAGTCTGCTGTTTCG 124
Db |||||
Qy 142 TCAGAGCGCCGAGCCAGGCTTCCCTCGGCTTCCGCTCGGAGAGCGCCGCGCGC 201
Db |||||
Qy 125 CGCGGCTTCCGCGCCACTCTCTCGCGCCAGAGTA-----GCCCGCGC 172
Db |||||
Qy 202 TGGCGAGCGGCTCATCAAGTCCCTTAACTCTGATGAAGAAGCAGCGGTGAAGCGC 261
Db |||||
Qy 173 TGGCGAAGGGCTGATCAAGTCCGCCAAGCCCTTAATGAAGAAGCAGCGGTGAAGCGC 232
Db |||||
Qy 262 ACCATCAAAACACCTTGGCAGCGCTTACAGTCTCGGCGCTCTGGAGACGCTGGGCAAGGCA 321
Db |||||
Qy 233 ACCACCAAGCACAACCTCGGCAACCGCTACAGTCTCTGGAGACCTTGGGCAAGGCA 292
Db |||||
Qy 322 CTTACGGGAAGGTGAAGAAGCAGCAGAGAGCTCGGCGCTCTGGTGGCCATCAAGTCCA 381
Db |||||
Qy 293 CTTACGGGAAGGTGAAGAAGCAGCAGAGAGCTCGGCGCTCTGGTGGCCATCAAGTCAA 352
Db |||||
Qy 382 TCAGGAAGCAAAATCAAAAGATGAGCAGGATCTGTGCAATATCGAGGGAGATTGAGA 441
Db |||||
Qy 353 TCCGGAAGCAAAATCAAAAGATGAGCAGGATCTGATGCAATACGAGGGAGATTGAGA 412
Db |||||
```



APPLICANT: RECIPON, Shirley A.  
APPLICANT: AZIMZAI, Yalda  
APPLICANT: POLICKY, Jennifer L.  
APPLICANT: DING, Li  
APPLICANT: GREYER, Megan  
APPLICANT: ELLIOTT, Vicki S.  
APPLICANT: THANGAVELU, Kavitha  
APPLICANT: BATRA, Sajeev  
APPLICANT: ISON, Craig H.  
TITLE OF INVENTION: HUMAN KINASES  
FILE REFERENCE: PI-0125 PCT  
CURRENT APPLICATION NUMBER: US/10/311,034  
CURRENT FILING DATE: 2002-12-10  
PRIOR APPLICATION NUMBER: 60/212,073; 60/213,467; 60/215,651; 60/216,605; 60/218,372;  
60/228,056  
PRIOR FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-07-25  
NUMBER OF SEQ ID NOS: 52  
SOFTWARE: PERL Program  
SEQ ID NO 38  
LENGTH: 3360  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No. US20040023242A1 4841542CB1  
US-10-311-034-38

Query Match 47.2%; Score 1370.6; DB 16; Length 3360;  
Best Local Similarity 81.5%; Pred. No. 0;  
Matches 1649; Conservative 0; Mismatches 339; Indels 35; Gaps 4;

82 CTACTGCTGCGCGACCACTCCACCTCGCGGTCCCGCCACCACCATGGAGTCGGTGCCTTAC 141  
39 CTACTGATTCCTCCGCGCCCTTCTCACTCTGCTCGGCATGAGTCGCTGTTTCG 98  
142 TCACGCGCGCGAGCAGCGCTCCCTCGCGCTCCCGCTCGGAGCGCCCGCGC 201  
99 CGCGCGCTCGCGCGCCACTCTCTCGCGC-----CAGAGCTAGCCCGCGC 145  
202 TGGCGGAGCGGCTCATCAAGTCGCTAACTCTGTGATGAGNAGCAGCGGTGAGCGGC 261  
146 TGGCGGAGGCGCTGATCAAGTCGCGCCCAAGCCCTTAATGAAGAGCAGCGGTGAAGCGC 205  
262 ACCATCAACAAACAACCTCGCGCAACCTGCGTACGAGTTCCTGGAGACGCTGGGCAAGGCA 321  
206 ACCACCAACAGCAACCTCGCGCAACCTGCGTACGAGTTCCTGGAGACCTGGGCAAGGCA 265  
322 CCTACGGGAAGGTGAAGAGCAAGAGCAGAGAGCTCGGGCGCTGTTGGCCATCAAGTCCA 381  
266 CCTACGGGAAGGTGAAGAGCGCGGAGAGCTCGGGCGCTGTTGGCCATCAAGTCAA 325  
382 TCAGGAAGCAAAATCAAGATGAGCAGGATCTGCTGCAATACGAGGAGATTGAGA 441  
326 TCGGAAGGACAAATCAAGATGAGCAGGATCTGATGCAATACGAGGAGATTGAGA 385  
442 TCATGCTTCACTCAACCAACCCACATCATTTGCCATCCATGAAGTGTGAGAAATAGCA 501  
386 TCATGCTCATCGCTCAACCAACCTCAATCATTCATTCATTCATGAGTGTGAGNAACAGCA 445  
502 GCAAGATTGTGATGTCATGAGTATGCCAGCGGAGCGGATCTGTATGATTTACATCAGTG 561  
446 GCAAGATCGTATCGTATGAGTATGCCAGCGGAGCGGATCTTTATGATATCATCAGCG 505  
562 AGCGGCCACCGCTGAGTGAAGGAGCGGAGCGGAGGATCTGCTGCAATACGAGGAGATTGAGA 621  
506 AGCGGAGAGAGCTCAGTGAAGCGGAGGAGGATCTGCTGCAATACGAGGAGATTGAGA 565  
622 TGCACTACTGCCACAGAAAGGAGTCTGTCACCGAGATCTCAAGCTGAGAAACATCTTC 681  
566 TGCACTATTGCCATCAGAAAGAGTGTTCACCGAGATCTCAAGCTGAGAAACATCTCT 625  
682 TAGATGCCAATGGAAACATCAAGATTGCTGACTTTGGCCTCTCCAACTGTACCAAAAG 741

626 TGGGTGCCAATGGGAATATCAAGATTGCTGACTTGGCCTCTCCAACTCTACCATCAAG 685  
742 GCAAGTTCTCCAGAGCTTCTGTGGGAGCCCTCTCTACGCTCGCCTGAGATAGTCAACG 801  
686 GCAAGTTCTCGAGACATTTCTGTGGGAGCCCTCTATGCTCTGCCAGAGATTGCAATG 745  
802 GGAAGCCCTATGTGGGCCAGAGGTGGACAGCTGGTCTCTGGGCGTCTCTCTGTACATCC 861  
746 GGAAGCCCTACACAGGCCAGAGGTGGACAGCTGGTCTCTGGGCTCTCTCTGTACATCC 805  
862 TGGTGCATGACACATGCTCTTTGACGGGAGGATCTAAACACACTGGTGAAGCAATCA 921  
806 TGGTGCATGACACATGCTCTTTGATGGGATGACCAATAAGATCTCTAGTGAACAGATCA 865  
922 GTAACGGGGCTTACCGTGAGCCGCCCAAGCCCTCGATGCTGTGGCTGATCCGGTGGC 981  
866 GCAACGGGGCTTACCGGAGCCACTTAACCTCTGATGCTGTGGCTGATCCGGTGGC 925  
982 TGTTAATGTGAACCCCGCTCGGCGCACTCTGGAGGATGTAGCCAGTCAATTTGGTGG 1041  
926 TGTGATGTGAACCCCGCTCGGCGCACTCTGGAGGATGTGGCCAGTCACTTTGGTGG 985  
1042 TCAACTGGGTTACACACCGGAGTGGGGAAACAGGAGCCCTGCTGTAGGCTGGGCAAC 1101  
986 TCAACTGGGCTACGCCACCCGAGTGGGAGAGCAGGAGGCTCCGCATGAGGCTGGGCAAC 1045  
1102 CTAGTGTGACTTTGGCGCGCTCCATGCGGAGCTGGTTACGCTGCTCTCGCGCCCC 1161  
1046 CTGGCAGTGACTTGGCGCGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1105  
1162 TCCTGGAGATGGAGCAAGGTGTGAGCTTTTCAAGAGCAGCAGTCCCGGGAGGTGGAA 1221  
1106 TCCTGGAGATGGGCGCAGGTGTGAGCTTTTCAAGAGCAGTGCATCTGCTGGGAGGAA 1165  
1222 GCACTGTACTTGGCTGGAGCGCAACATTTCTTTAAGAGTCCCGAAGAGAGATGACA 1281  
1166 GCACCACTTGGCTGGAGCGCAGCATTCGCTCAAGAGTCCCGAAGAGATGACA 1225  
1282 TGCTCAAAATCTGCAAGGTGACCCGCTGAGGATACCTCTTCCGCTCGGCGGAGAGCA 1341  
1226 TGCCCACTCTCTCAAGTGAACGCTGATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1285  
1342 GCCTTAAGCTTCCGAAAGGCAATTTCTCAAGAAAGTCTCTTACCTCTGCTGAGGGAGGTAC 1401  
1286 ACCTCAAGCTGCCAAGGGCAATTTCTCAAGAGAGGTGTGAGCTCTGCTGAGAGGGGTAC 1345  
1402 AGAGAGCTCTCAGGAACTCAGACCGGTGCTGATTAATCTCAGGGAGCCTGCTGCTGCTG 1461  
1346 AGAGAGCTCTCAGGAGCTCAGCCCAATCTCTGAGCGCCAGGGAGGCTGCTGCTGCTGCTG 1399  
1462 TATCCCTGCTCCCAAGGAGGATCTCTTAAGAGTCTCCAGAGCTGAGTGTGTTGTTACT 1521  
1400 ---CCCTGCTCCCAAGAGGGCAATTTCTCAAGAGGCGCCAGCAGCGGAGTCTGCTGCT 1456  
1522 ACTCTCTCCAGAGCCAGCGAGTCTGGGAACTTTAGACGCGCAGTGTGTTGTTGTA 1581  
1457 ACTCTCTCCAGAGCCAGTGAATCTGGGAGCTCTTGGAGCGCAGGAGCGTGTGTTGTA 1516  
1582 GTGGGAGCCCGTGGAGAGAGAGTCTCCACAGGCTTCCAGGGCTCTCTCTCCACCGCAAGG 1641  
1517 GTGGGAGTCCCAAGAGAGAGAGCTCCGCAAGCTTCCAGGGCTGCTCTCTCATCGCAAG 1576  
1642 GCATTTCTCAAACTCAATGGCAAGTCTCTCCAGAGCAGCTTTGGAGCTCGCGGCCCCACCA 1701  
1577 GCATTTCTCAAACTCAATGGCAAGTCTCTCCAGAGCAGCTTTGGAGCTCGCGGCCCCACCA 1636  
1702 CTTTGGCTCCCTGGAGCAACTGCTGCTCTCCATCTGAGCGCCCGGCGCGGCGCT 1761  
1637 CTTTGGCTCCCTGGAGCAACTGCTGCTCTCCATCTGAGCGCCCGGCGCGGCGCT 1696  
1762 CAGGGCTGTGAGTGGAGCAGCATCTCTGCTCCGAGTCTTTGACCAATGAGCTTGC 1821

Db 1697 CAGGGCTGTGAGGAGGACAGCATCTCTGTCTCTAGTCTCTTGTAGCCAGCTGACTTGC 1756  
Qy 1822 CTGAAGCTTTCCCGAACCCTGAGGGGCTGTGTCTGTGAGCAACCTGAGGGGC 1881  
Db 1757 CTGAAGCTTTCCCGAACCCTGAGGGGCTGTGTCTGTGAGCAACCTGAGGGGC 1816  
Qy 1882 TTGAGCAGCTTCCCTCAGAAG-----GTCTGAAGCGATGGTGGCAGGAATCCT 1929  
Db 1817 TTGAGGAGCCCTCTGAGAGGCTTGAAGCTGCTGAGGCGCTGGCGGAGAGTCTT 1876  
Qy 1930 TGGGGGATAGTCTTTTCTGTGACAGCTGCCAAGAGGTGATGTCAGCCTTACAGACAAG 1989  
Db 1877 TGGGGGATAGTCTTTTCTGTGACAGCTGCCAAGAGGTGATGTCAGCCTTACAGACAAG 1936  
Qy 1990 CCTAGGAATCTGCTCAAGCTCAGCTGAGGAGGAGATGGTCCCTAGTATGGGTAG 2049  
Db 1937 CACTGAGGCTGTCTCAAGCTCACTGAGTGGAGTAGGCAATGCCCGAG-CCCGGTGAG 1995  
Qy 2050 GCTCTGAGAGGTTTTCAGAGAGAACCTCTGGTGGGATTCCTCC 2092  
Db 1996 GCTCTGAGATGACGTGTTGTCACCCCGAGGGAGATGCTTTC 2038

## RESULT 14

US-09-963-159-3  
; Sequence 3, Application US/09963159  
; Patent No. US2002007312A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Galvin, Katherine M.  
; TITLE OF INVENTION: 3700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR  
; FILE REFERENCE: 10147-50U1  
; CURRENT APPLICATION NUMBER: US/09/963.159  
; CURRENT FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 60/234,922  
; PRIOR FILING DATE: 2000-09-25  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 1884  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-963-159-3

Query Match 46.9%; Score 1360.2; DB 9; Length 1884;  
Best Local Similarity 84.2%; Pred. No. 0;  
Matches 1565; Conservative 0; Mismatches 273; Indels 21; Gaps 2;

Qy 169 CCTCCGCTGCTCGGAGAGCGCCCGCGCTGCGGAGCGGCTCATCAAGTCGCCTA 228  
Db 35 CCACTCCCTCGGCGCAGAGCTAGCCCGCGCTGCGGAGGCTGATCAAGTCGCCCA 94  
Qy 229 AACTCTGATCAAGACAGCGGTGAAGCGGACCATCAAAACAACTGCGGCACC 288  
Db 95 AGCCCTTAATGAAGACAGCGGTGAAGCGGACCATCAAAACAACTGCGGCACC 154  
Qy 289 GCTACAGATTCCTGGAGACGCTGGGCAAGGCGACCTACGGGAGGTGAAGAGCGAG 348  
Db 155 GCTACAGATTCCTGGAGACGCTGGGCAAGGCGACCTACGGGAGGTGAAGAGCGAG 214  
Qy 349 AGAGCTCGGGCGCTGCTGGTGGCCATCAAGTCCATCAGGAAGACAAATCAAAATGAGC 408  
Db 215 AGAGCTCGGGCGCTGCTGGTGGCCATCAAGTCCATCAGGAAGACAAATCAAAATGAGC 274  
Qy 409 AGAGCTCGGCTGACATACGAGGAGATTCAGATCATGCTTCACTCAACCAACCCCA 468  
Db 275 AGATCTGATGACATACGAGGAGATTCAGATCATGCTTCACTCAACCAACCCCA 334  
Qy 469 TCATTGCCATCCATGAAGTGTGAGAAATAGCAAGATTTGATGTCATGAGATG 528  
Db 335 TCATTGCCATCCATGAAGTGTGAGAAATAGCAAGATTCGATGTCATGAGATG 394  
Qy 529 CCAGCCGAGGCGATCTGTATGATTATCATAGTAGCGGCCACCGCTGAGTGAAGCGGAGC 588

Db 395 CCAGCGGGCGACCTTTATGACTACATCAGCAGCGGAGCAGCTCAGTGAAGCGCGAG 454  
Qy 589 CAGGCAATTTCTTCGACAGATCGTGTCTGCCCTGCACTACTGCCACAGACGGGATCG 648  
Db 455 CTAGGCAATTTCTTCGCGAGATCGTCTCTGCGGTGCACTATTGCCATCAGAACAGATTG 514  
Qy 649 TTCAACGAGATCTCAAGCTGGAACATCCTTCTAGATGCCAATGGAAACATCAAGATTG 708  
Db 515 TCACCGAGATCTCAAGCTGGAACATCCTTCTGGATGCCAATGGAAATATCAAGATTG 574  
Qy 709 CTGACTTTGGCCCTCTCAACCTGTACCAAAAGCAAGTCTCTCCAGACGTTCTGTGGGA 768  
Db 575 CTGACTTTGGCCCTCTCAACCTGTACCAATCAAGCAAGTCTCTGCGACATTTCTGTGGGA 634  
Qy 769 GCCTCTCTACGGCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCGAGGTGG 828  
Db 635 GCCTCTCTATGCTGCGCAGAGATTTCTCAATGGGAAGCCCTACACAGGCCCGAGGTGG 694  
Qy 829 ACAGCTGCTCTGCGGCTTCTCTGTACATCCTGTGTGATGCGACCATGCTCCCTTTGACG 888  
Db 695 ACAGCTGCTCTGCGGCTTCTCTGTACATCCTGTGTGATGCGACCATGCTCCCTTTGATG 754  
Qy 889 GGCAGGATCAATAAACAACCTGCTGAAGCAATCAGTAACGGGGCTTACCGTGAAGCGGCCA 948  
Db 755 GGCATGACCATAGATCTTAGTGAACAGATCAGCAACGGGGCTTACCGGAGGCCACTA 814  
Qy 949 AGCGTCCGATGCTGTGGCTGATCGGTGGCTGTGTAATGTGTGAACCCACCGCTCGG 1008  
Db 815 AACCTCTGATGCTGTGGCTGATCGGTGGCTGTGTAATGTGTGAACCCACCGCTCGG 874  
Qy 1009 CCACACTGGAGGATGAGCCAGTCACTGTGGTCACTGGGCTTACACCAACCGAGTTCG 1068  
Db 875 CCACCTGGAGGATGAGCCAGTCACTGTGGTCACTGGGCTTACACCAACCGAGTTCG 934  
Qy 1069 GGGAAACAGGAAGCCCTGCTGAGGGTGGGCAACCTAGTGTGTGCTTTGGCGGGCTTCCA 1128  
Db 935 GAGAGCAGGAGGCTCCGCTAGGGTGGGCAACCTGCGAGTGTCTGTGCCCGGCTTCCA 994  
Qy 1129 TGGCGACTGCTTACGTGCTCTCGCGGCTTCTCGCGGCTTCTCGCGGCTTCTCGCGGCTTCCA 1188  
Db 995 TGGCTGACTGCTGCTCGCGGCTTCTCGCGGCTTCTCGCGGCTTCTCGCGGCTTCCA 1054  
Qy 1189 GCTTCTTCAAGCAGACGCTGCGGGAGGTGGAAGCACTGTACCTGGCTGAGCGGCAAC 1248  
Db 1055 GCTTCTTCAAGCAGACGCTGCTGCTGGGGAGGCAACCCCTGCGCTGAGCGGCGAGC 1114  
Qy 1249 ATTCTTTAAGAGTCCGAAAGAGGAGATGACATGGCTCAAAATCTGCAAGGTGACCGCG 1308  
Db 1115 ATTCTGCTCAAGAGTCCGCAAGGAGATGACATGGCTTCTCTCCACAGTGACAGCG 1174  
Qy 1309 CTGAGATACCTTCTTCTCGGCTTGGCAAGGAGGCTTAAAGTCTCGAAAGGATTTCTCA 1368  
Db 1175 CTGATGACACTGCTTCTCGGCTTGGCAAGGAGGCTTAAAGTCTCGAAAGGATTTCTCA 1234  
Qy 1369 AGAAAGTCTCTTCTGCTGAGGAGGAGTACAGGAGGAGGCTTAAAGTCTCGAAAGGATTTCTCA 1428  
Db 1235 AGAAAGTCTCTTCTGCTGAGGAGGAGTACAGGAGGAGGCTTAAAGTCTCGAAAGGATTTCTCA 1294  
Qy 1429 TGCTGATCTTCAAGGAGGAGGAGTCTTCTGCTGATGATGCTTCTCGAAAGGAGGATTTCTCA 1488  
Db 1295 TCCCTGAGGCTTCAAGGAGGAGGAGTCTTCTGCTGATGATGCTTCTCGAAAGGAGGATTTCTCA 1345  
Qy 1489 TTAAGAGTCTGAGAGGAGGAGTCTTCTGCTGATGATGCTTCTCGAAAGGAGGATTTCTCA 1548  
Db 1346 TCAAGAGGCTTCAAGGAGGAGGAGTCTTCTGCTGATGATGCTTCTCGAAAGGAGGATTTCTCA 1405  
Qy 1549 GGGAACTCTTCAAGGAGGAGGAGTCTTCTGCTGATGATGCTTCTCGAAAGGAGGATTTCTCA 1608  
Db 1406 GGGAGCTCTTCAAGGAGGAGGAGTCTTCTGCTGATGATGCTTCTCGAAAGGAGGATTTCTCA 1465  
Qy 1609 CACAGGCTTCAAGGAGGAGGAGTCTTCTGCTGATGATGCTTCTCGAAAGGAGGATTTCTCA 1668

Db 1466 CGCAAGCTTCAGGCTCCTCTCCATCGCAAGGATCTCTCAAACTCAATGGCAAGTTCT 1525  
Qy 1669 CCCGACAGCTTAGAAGGCACTACCCCTAGCACCTTTGGCTCCCTGAGCAACAACTGGCCT 1728  
Db 1526 CCAGACAGCTTAGAGCTGGGGCCCCCACCACCTTCGGCTCCTGGATGAATCGCCC 1585  
Qy 1729 CTTCCATCTCGAGCCGCGCCAGCCCTCAGGGGCTGTGAGTGAGGACAGCATCC 1788  
Db 1586 CACCTCGCCCTCGCCGCGCCAGCCGAGCCTCAGGGGCTGTGAGCGAGGACAGCATCC 1645  
Qy 1789 TGTCTCCGAGCTCTTTGACCAATGCACTGCTGCACTCTTCCGAAACCCCACTGA 1848  
Db 1646 TGTCTCTGAGTCTTTTACAGAGTGGACTTGCCTGAACGGCTCCAGAGCCGCCACTGC 1705  
Qy 1849 GGGGCTGTGTCTGTGCAACACCTGAGGGGCTTTGAGCAGCCTCCCTCAGAAAG- 1902  
Db 1706 GGGGCTGTGTCTGTGCAACACCTCAGGGGCTTTGAGGAGCCCTCAGAGGGCCCTG 1765  
Qy 1903 -----GTCTGAAGCGATGTGGGAGGAATCTTTGGGGGATAGTGTCTTCTGACAG 1956  
Db 1766 GAAGCTGCTGAGCGCTGGCGGAGGATCTTTGGGGGACAGCTGCTTTTCCCTGACAG 1825  
Qy 1957 ACTGCCAAGAGTGAAGTGCAGCCTTACAGACAGCCCTAGGAATCTGCTCAAGCTCAGC 2015  
Db 1826 ACTGCCAAGAGTGAAGGAGCAGCCTTACCGACAGGCACTGAGGGTCTGCTCAAGCTCACC 1884

RESULT 15  
US-10-423-543-45  
; Sequence 45, Application US/10423543  
; Publication No. US20040058355A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Libermann, Rosana K.  
; APPLICANT: Hunter, John J.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Rudolph-Owen, Laura A.  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Tsai, Fong-Ying  
; APPLICANT: Galvin, Katherine M.  
; APPLICANT: Chun, Miyoung  
; APPLICANT: Williamson, Mark J.  
; APPLICANT: Silos-Santiago, Immaculada  
; APPLICANT: Bandaru, Rajasekhar  
; TITLE OF INVENTION: NOVEL 21910, 56634, 55053, 2504, 15977,  
; TITLE OF INVENTION: 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638,  
; TITLE OF INVENTION: 18610, 33217, 21967, h1983, m1983, 38555 OR 593 MOLECULES  
; TITLE OF INVENTION: AND USES THEREFOR  
; FILE REFERENCE: MPI03-0230MIM  
; CURRENT APPLICATION NUMBER: US/10/423,543  
; CURRENT FILING DATE: 2003-04-25  
; PRIOR APPLICATION NUMBER: US 10/278,036  
; PRIOR FILING DATE: 2002-10-22  
; PRIOR APPLICATION NUMBER: US 09/711,216  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: US 60/205,447  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 10/012,055  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/248,325  
; PRIOR FILING DATE: 2000-11-14  
; PRIOR APPLICATION NUMBER: US 10/003,690  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/248,893  
; PRIOR FILING DATE: 2000-11-15  
; PRIOR APPLICATION NUMBER: US 09/797,039  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/186,061  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 10/217,168  
; PRIOR FILING DATE: 2002-08-12  
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45  
; LENGTH: 1884  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-10-423-543-45  
Query Match 46.9%; Score 1360.2; DB 16; Length 1884;  
Best Local Similarity 84.2%; Pred. No. 0;  
Matches 1565; Conservative 0; Mismatches 273; Indels 21; Gaps 2;  
Qy 169 CTTCCGCTTGGCTTCGAGAGCCCGCGCTGCGGAGCGGCTCATCAAGTCGCTA 228  
Db 35 CCACTCCCTCGCGCGCAGAGCTAGCCGCGCTGGCGAAGGCTGATCAAGTCGCCA 94  
Qy 229 AACCTCTGATGAAGAAGCAGCGGCTGAAGCGGCACCATCAAAACACAACTCGGGCACC 288  
Db 95 AGCCCCCTAATGAAGAAGCAGCGGCTGAAGCGGCACCAACCAAGCAACCTCGGGCACC 154  
Qy 289 GCTACAGATTTCTTGGAGAGCTGGGCAAGGCGACCTACGGAAGGTTGAAGAAGCGCAG 348  
Db 155 GCTACAGATTTCTTGGAGAGCCCTGGGCAAGGCGACCTACGGAAGGTTGAAGAAGCGCGG 214  
Qy 349 AGAGCTCGGGCGCTCTGGTGGCCATCAAGTCCATCAGGAAGACAAATCAAAAGATGAGC 408  
Db 215 AGAGCTCGGGCGCTTGGTGGCCATCAAGTCCATCAGGAAGACAAATCAAAAGATGAGC 274  
Qy 409 AGGATCTGTGCACATACGAGGAGATTTGAGATCATGTCTTCACTCAACACCCCA 468  
Db 275 AAGATCTGATGCACATACGAGGAGATTTGAGATCATGTCTTCACTCAACACCCCTCACA 334  
Qy 469 TCATTGCCATCCATGAAGTGTGTTGAGAATAGCAGCAAGATTGTTGATTGTTCATGGAGTATG 528  
Db 335 TCATTGCCATCCATGAAGTGTGTTGAGAATAGCAGCAAGATTGTTGATTGTTCATGGAGTATG 394  
Qy 529 CCAGCCGAGCGATCTGTATGATTATCATCAGTGAGCGGCCACCGCTGAGTGAGCGGACG 588  
Db 395 CCAGCCGAGCGATCTGTATGATTATCATCAGTGAGCGGCCACCGCTGAGTGAGCGGACG 454  
Qy 589 CAGGCAATTTCTCCGACAGATCGTGTGCGCTGCACTACTGCGCCACGAGCGGATCG 648  
Db 455 CTAGGCAATTTCTCCGCGAGATCGTCTCTGCGCTGCACTATTGCGCATCAGAACAGAGTTG 514  
Qy 649 TTCACCGAGATCTCAAGCTGGAAACATCTCTTACATGCAATGCAATGAAACATCAAGATTG 708  
Db 515 TCACCGAGATCTCAAGCTGGAGACATCTCTTGGATGCAATGGGAATATCAAGATTG 574  
Qy 709 CTGACTTTGGCTCTCCAACTGTACCAAAAGGCAAGTTCTCTCCAGACGTTCTGTGGGA 768  
Db 575 CTGACTTTGGCTCTCCAACTGTACCAAAAGGCAAGTTCTCTCCAGACATTTCTGTGGGA 634  
Qy 769 GCGCTCTACGCTCGCTGAGATAGTCAAGGGAAGCCCTATGTGGGCCACAGAGTTG 828  
Db 635 GCGCTCTATGCTCGCCAGAGATTGTCAATGGGAAGCCCTACACAGGCGCCAGAGTTG 694  
Qy 829 ACAGCTGCTCTGCGGCTTCTCTGTACATCTCTGTGATGGCAGCACCATCCCTTTTGACG 888  
Db 695 ACAGCTGCTCTGCGGCTTCTCTGTACATCTCTGTGATGGCAGCACCATCCCTTTTGATG 754  
Qy 889 GGCAGGATCAAAACACATGTTGGAAGCAAAATCAAGTAAAGCGGGCTTACCGTGAGCGGCCA 948  
Db 755 GGCATGACCATAGATCTTAGTGAACAGATCAGAACGCGGGCTACCGGGAGCCACTA 814  
Qy 949 AGCGCTCCGATGCTGTGGCTGATCCGGTGGCTGTTAATGTTGAACCCCAACCGCTCGG 1008  
Db 815 AACCTCTGATGCTGTGGCTGATCCGGTGGCTGTTGATGTTGTTGAACCCCAACCGCTCGG 874  
Qy 1009 CCACACTGGAGGATGTAGCCAGTCAATTGTTGGTCAACTGGGTTTACACACCGGAGTCG 1068  
Db 875 CCACCTGGAGGATGTGGCAGTCACTGTTGGTCAACTGGGCTAGCCACCCCGAGTGG 934  
Qy 1069 GGAACAGGAAGCCCTGCGTGAGGCTGGGCAACCTAGTGTGACTTTTGGCGGGCTCCA 1128

Db 935 GAGAGCAGAGGCTCCGATGAGGCTGGGCAACCTGGCAGTGACTCTGCCCGCGCTCCA 994  
Qy 1129 TGGCGGACTGTTAGTCTGCTCTGCGCCCGCTCTGGAGAATGGAGCCAAAGGTGTGCA 1188  
Db 995 TGGCTGACTGGCTCGCGGCTTCTCCCGCCCTCTCTGGAGATGGGGCCAAAGGTGTGCA 1054  
Qy 1189 GCTTCTTCAAGCAGCAGCTGCGGGAGTGAAGGACTGTACTTGGGCTGGAGCGGCAAC 1248  
Db 1055 GCTTCTTCAAGCAGCAGCTGCTGGTGGGGAGGACCAACCCCTGGCCCTGGAGCGCCAGC 1114  
Qy 1249 ATTCTTTAAGAAAGTCCGAAAGAGAGATGACATGGCTCAAAATCTGCAAGGTGACCCGG 1308  
Db 1115 ATTGCTCAAGAGTCCCGCAAGGAGATGACATGGCCAGTCTCTCCAAGTGACACGG 1174  
Qy 1309 CTGAGGATACCTTTCTGCGCCTGGCAAGAGCAGCTTTAAGCTTCCGAAAGGCATTCTCA 1368  
Db 1175 CTGATGACACTGCCCCATCGCCCTGGCAAGAGCAACCTCAAGCTGCCAAAGGGCATTTCTCA 1234  
Qy 1369 AGAAAAAGTCTCTACCTCGTCAAGGAGGTACAGGAGGACCTCAGGAATCAGACGG 1428  
Db 1235 AGAAGAGGTGTAGCCTTCTGCAAGGGGTACAGGAGGACCTTCCGAGGCTCAGCCCAA 1294  
Qy 1429 TGCCTGATATCTCAGGGCAGCTGCTCCTGTGTATCCCTGCTCCCAAGGAAAGGCATCC 1488  
Db 1295 TCCCTGCGAGCCAGGGCAGCTGCC-----CCCTGCTCCCCAAGAGGGCATTC 1345  
Qy 1489 TTAGAAGTCTCGACAGGTGAATCTGGTTACTACTCTCTCCAGAGCCAGCGAGTCTG 1548  
Db 1346 TCAAGAGCCCGCAGCAGCGAGTCTGGCTACTACTCTCTCCCGAGCCAGTGAATCTG 1405  
Qy 1549 GGGAACTCTTAGACGCGCAGTATGTTGTGTGGGGACCCGCTGGAGCAGAGTCTC 1608  
Db 1406 GGGAGCTCTTGGAGCGAGCGACGTGTTGTGTGGGATCCCAAGAGCGAGGCTC 1465  
Qy 1609 CACAGGCTTACGGCTCCTCTCCACCGCAAGGGCAATTTCAAACTCAATGGCAAGTTCT 1668  
Db 1466 CGCAAGCTTCAGGGCTCTCTCTCCATCGCAAGGCATCTCAAACTCAATGGCAAGTTCT 1525  
Qy 1669 CCGCAGAGCTTAGAAGGCACTACCCCTAGCAGCTTTGGCTCCCTGGAAGCACTGGCCT 1728  
Db 1526 CCCAGACAGCCTTAGAGCTCGCGGCCCGCCACCTTCGGCTCCCTGGATGAATCGCCC 1585  
Qy 1729 CCTCCCATCTGACGCGCGCCCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCC 1788  
Db 1586 CACCTCGCCCTCGCCCGCGGCGAGCCGACCTCAGGGGCTGTGAGCGAGGACAGCATCC 1645  
Qy 1789 TGTCTCCGAGTCTTTTGAACAATTGCACTTGGCTGAACGCTTTCCCGAAACCCCACTGA 1848  
Db 1646 TGTCTCTGAGTCTTTTGAACAAGTGGACTTGGCTGAACGCTTCCAGAGCCCCCACTGC 1705  
Qy 1849 GGGGCTGTGTGTGTGTCGACCACTGAGGGGCTTGAAGCAGCTCCCTCTAGAAG----- 1902  
Db 1706 GGGGCTGTGTGTGTGTCGACCACTCAGGGGCTTGAAGGAGCCCCCTCAGAGGGGCCCTG 1765  
Qy 1903 -----GTCTGAAGCGATGTGGCAGGAATCTTTGGGGGATAGCTGCTTTTCTGACAG 1956  
Db 1766 GAAGCTCCCTGAGGGGCTGGGGCAGGATCCTTTGGGGGACAGCTGCTTTTCTGACAG 1825  
Qy 1957 ACTGCCAAGAGGTGATGTCAGCTTACAGAGCCCTTAGGAATCTGCTCAAAGCTCAGC 2015  
Db 1826 ACTGCCAAGAGGTGACAGCGACCTTACCAGAGGCACTGAGGGTCTGCTCAAAGCTCACC 1884

Search completed: January 25, 2005, 07:55:03  
Job time : 1528.67 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 14:50:33 ; Search time 12179.3 Seconds  
(without alignments)  
11267.845 Million cell updates/sec

Title: US-09-980-464-4  
Perfect score: 2902  
Sequence: 1 cactagtggatccaagaat.....aaaaaaaaaaaaaaaaaaaaa 2302

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2902	100.0	2902	6	AR279568
2	2820.4	97.2	2938	10	BC046833 Mus muscu
3	2763.8	95.2	2917	10	BC033302 Mus muscu
4	2066.8	71.2	2929	6	AX380960 Sequence
5	2022.2	69.7	2027	6	AX381046 Sequence
6	1771.4	61.0	2026	6	AX380986 Sequence
7	1593.2	54.9	291762	2	AC125887 Rattus no
8	1379	47.5	3353	6	AX407116 Sequence
9	1375.8	47.4	3395	6	CQ782778 Sequence
10	1375.8	47.4	3395	6	BD127309 Primer fo
11	1375.8	47.4	3395	9	AK074830 Homo sapi
12	1375.8	47.4	3402	9	BC017306 Homo sapi
13	1374.2	47.4	3443	6	AX086949 Sequence
14	1374.2	47.4	3443	6	AX780351 Sequence
15	1374.2	47.4	3443	9	AX780351 Sequence
16	1370.6	47.2	3360	6	AX642961 Homo sapi
17	1360.2	46.9	1884	6	AX407118 Sequence
18	1360	46.9	1887	6	CQ714327 Sequence
19	1349.2	46.5	2291	6	AX056394 Sequence

c	20	1278.6	44.1	252978	2	AC126983	AC126983 Rattus no
	21	1230	42.4	1833	6	AX803413	AX803413 Sequence
	22	1071.4	36.9	1186	6	AX380972	AX380972 Sequence
	23	943.4	32.5	1186	6	AX381029	AX381029 Sequence
	24	849.2	29.3	2616	6	AX430369	AX430369 Sequence
	25	849.2	29.3	2619	6	AX430468	AX430468 Sequence
c	26	732	25.2	179414	9	AC098933	AC098933 Homo sapi
	27	728.8	25.1	6389	9	HSM06181	BSX37564 Homo sapi
	28	675.4	23.3	741	6	AX380991	AX380991 Sequence
	29	651.8	22.5	1274	5	BX950299	AX380991 Gallus ga
	30	648	22.3	707	6	AX381050	AX381050 Sequence
	31	648	22.3	712	6	AX380990	AX380990 Sequence
	32	644.2	22.2	661	6	AX381052	AX381052 Sequence
	33	643	22.2	734	6	AX381051	AX381051 Sequence
	34	583.2	20.1	588	6	AX381056	AX381056 Sequence
	35	557.4	19.2	660	6	AX380992	AX380992 Sequence
	36	534	18.4	578	6	AX381054	AX381054 Sequence
	37	515	17.7	548	6	AX381055	AX381055 Sequence
	38	513	17.7	521	6	AX381053	AX381053 Sequence
	39	511.6	17.6	579	6	AX380994	AX380994 Sequence
	40	507.6	17.5	6825	6	CQ715220	CQ715220 Sequence
	41	507.6	17.5	6828	6	CQ776374	CQ776374 Sequence
	42	507.6	17.5	6828	6	AX213156	AX213156 Sequence
	43	507.6	17.5	6828	9	AB011109	AB011109 Homo sapi
	44	506.4	17.5	1986	6	AX213158	AX213158 Sequence
	45	506.4	17.5	2884	6	AX213160	AX213160 Sequence

ALIGNMENTS

RESULT 1  
AR279568  
LOCUS AR279568 2902 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 4 from patent US 6514719.  
ACCESSION AR279568  
VERSION AR279568.1 GI:29714427  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2902)  
AUTHORS Bird,T.A., Virca,G.D., Martin,U. and Anderson,D.M.  
TITLE Methods for identifying compounds that alter kinase activity  
JOURNAL Patent: US 6514719-A 4 04-FEB-2003;  
FEATURES Location/Qualifiers  
source 1..2902  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN	Query Match	100.0%;	Score 2902;	DB 6;	Length 2902;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2902;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CAC TAGTGGATCCAAAGAAATTCGGCAGCGCGTGTCTCGGGTGGCGGTGTGACCTCTGAG	60		
Db	1	CAC TAGTGGATCCAAAGAAATTCGGCAGCGCGTGTCTCGGGTGGCGGTGTGACCTCTGAG	60		
Qy	61	CCCGCGGCTCAGCGCGGCTGTCTGTGTCGCCGACCCACTCACTCTCGGGGTCCCGGCA	120		
Db	61	CCCGCGGCTCAGCGCGGCTGTCTGTGTCGCCGACCCACTCACTCTCGGGGTCCCGGCA	120		
Qy	121	CCATGAGTGGTGGCTTACTCCAGCGCCGAGGCTCCCTCGGCTCCGCGCTCG	180		
Db	121	CCATGAGTGGTGGCTTACTCCAGCGCCGAGGCTCCCTCGGCTCCGCGCTCG	180		
Qy	181	CCTCGGAGAGCGCCGCGCGCTGTCTCAAGTCGGCTCAAGTCCCTAAACCTCTGATGA	240		
Db	181	CCTCGGAGAGCGCCGCGCGCTGTCTCAAGTCGGCTCAAGTCCCTAAACCTCTGATGA	240		
Qy	241	AGAAGCAGCGGTGAAGCGCGCACCATCACAAAACAACTGCGGACCGCTAGAGTTCC	300		







Best Local Similarity 99.1%; Pred. No. 0;			
Matches 2855; Conservative 0; Mismatches 1; Indels 24; Gaps 1;			
Qy	43	GGGCTGTGACCTCTGAGCCCGCGCTCAGCGCGCTGCTACTCTGCTCCGAGCCACTC	102
Db	1	CGGCTGTGACCTCTGAGCCCGCGCTCAGCGCGCTGCTACTCTGCTCCGAGCCACTC	60
Qy	103	CACCTCGGGTCCCGGACCATGAGTCGGTGGCTTACTCAGGGCCGAGCCAGGCTC	162
Db	61	CACCTCGGGTCCCGGACCATGAGTCGGTGGCTTACTCAGGGCCGAGCCAGGCTC	120
Qy	163	CCTCGGCTCCGCGCTCGAGAGCGCCCGCGCTGCGGAGCGGCTCATCAAGT	222
Db	121	CCTCGGCTCCGCGCTCGAGAGCGCCCGCGCTGCGGAGCGGCTCATCAAGT	180
Qy	223	CGCTTAACCTCTGATGAAGAGCAGCGGTGAAGCGGCACCATCACAAACACACCTGC	282
Db	181	CGCTTAACCTCTGATGAAGAGCAGCGGTGAAGCGGCACCATCACAAACACACCTGC	240
Qy	283	GGCACCCTACGAGTCTCTGAGAGCGCTGGGCAAGGCACTACGGGAGGTGAAGAGG	342
Db	241	GGCACCCTACGAGTCTCTGAGAGCGCTGGGCAAGGCACTACGGGAGGTGAAGAGG	300
Qy	343	CACGAGAGCTCGGGCGCTCTGCTGGCCATCAAGTCCATCAGGAAGACAAAATCAAAG	402
Db	301	CACGAGAGCTCGGGCGCTCTGCTGGCCATCAAGTCCATCAGGAAGACAAAATCAAAG	360
Qy	403	ATGAGCAGGATCTGCTGACATACGAGGGAGATTGAGATCATGTCTCACTCAACACC	462
Db	361	ATGAGCAGGATCTGCTGACATACGAGGGAGATTGAGATCATGTCTCACTCAACACC	420
Qy	463	CCCATCATTTGCCATCCATGAA-----GTGTTTCAGAA	498
Db	421	CCCATCATTTGCCATCCATGAGTGGGAGATCACGCTTAGTGACTGTGTTGAGAATA	480
Qy	499	GCAGCAAGATTGATTTGTATGAGTATGCCAGCGGAGTCTGTATGATTACATCA	558
Db	481	GCAGCAAGATTGATTTGTATGAGTATGCCAGCGGAGTCTGTATGATTACATCA	540
Qy	559	GTGAGCGGCCACGGCTGAGTGGCGGACGCGAGGCAATTTCTCGGACAGATCGTGTG	618
Db	541	GTGAGCGGCCACGGCTGAGTGGCGGACGCGAGGCAATTTCTCGGACAGATCGTGTG	600
Qy	619	CCCTGCACTACTGCCACAGAAACGGGATCGTTTACCGAGATCTCAAGCTGGAAACATCC	678
Db	601	CCCTGCACTACTGCCACAGAAACGGGATCGTTTACCGAGATCTCAAGCTGGAAACATCC	660
Qy	679	TTCTAGATGCCAATGGAAACATCAAGATTGCTGACTTTGGCTCTCCAACTGTACCA	738
Db	661	TTCTAGATGCCAATGGAAACATCAAGATTGCTGACTTTGGCTCTCCAACTGTACCA	720
Qy	739	AAGCAAGTTTCTCCAGACGTTCTGTGGAGCCCTCTCTAGCCCTCGCCTGAGATGCA	798
Db	721	AAGCAAGTTTCTCCAGACGTTCTGTGGAGCCCTCTCTAGCCCTCGCCTGAGATGCA	780
Qy	799	ACGGAAACGCTTATGTGGGCGCCAGAGGTGGACAGCTGCTCTGGGGGTTCTCTGTACA	858
Db	781	ACGGAAACGCTTATGTGGGCGCCAGAGGTGGACAGCTGCTCTGGGGGTTCTCTGTACA	840
Qy	859	TCCTGTGATGGACCATGCTTTGACGGGAGGATCATATAAACAATGGTGAAGCAAA	918
Db	841	TCCTGTGATGGACCATGCTTTGACGGGAGGATCATATAAACAATGGTGAAGCAAA	900
Qy	919	TCAGTAACGGGGTTACCGTGGCGCCCAAGCGCTCCGATGCTGTGGCTCATCCGT	978
Db	901	TCAGTAACGGGGTTACCGTGGCGCCCAAGCGCTCCGATGCTGTGGCTCATCCGT	960
Qy	979	GGCTGTAAATGGTGAACCCCAACCGCTGGGCCACATCGAGGATGTAGCCAGTCAITGGT	1038
Db	961	GGCTGTAAATGGTGAACCCCAACCGCTGGGCCACATCGAGGATGTAGCCAGTCAITGGT	1020
Qy	1039	GGCTCACTGGGTTACACACCGGATCGGGNACAGGAAGCCCTGCGTGGGGTGGC	1098

1021	GGGTCAACTGGGTTACACACCGAGTCGGGGAACAGGAAGCCCTCGGTAGGGTGGGC	1080
1099	ACCTAGTGTGACTTTGGCGCGGCTCCATAGCGGACTGGTTACGTGCTCTCTCGCGCC	1158
1081	ACCTAGTGTGACTTTGGCGCGGCTCCATAGCGGACTGGTTACGTGCTCTCTCGCGCC	1140
1159	CCCTCTGGAGATGGAGCCAAAGTGTGAGCTTTCTTAAGAGCAGCAGTGGCGGGAGTG	1218
1141	CCCTCTGGAGATGGAGCCAAAGTGTGAGCTTTCTTAAGAGCAGCAGTGGCGGGAGTG	1200
1219	GAAGCACTGTACCTGGCTGAGCGGCAATCTCTTAAGAGTCCCGAAAGAGAGATG	1278
1201	GAAGCACTGTACCTGGCTGAGCGGCAATCTCTCTTAAGAGTCCCGAAAGAGAGATG	1260
1279	ACATGGCTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCTCTTCTGCGCCCTGGCAAGA	1338
1261	ACATGGCTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCTCTTCTGCGCCCTGGCAAGA	1320
1339	GCAGCCTTAAGCTTCGAAAGGCAATCTCAAGAAAAGTCTCTACCTGCTCAGGGAGG	1398
1321	GCAGCCTTAAGCTTCGAAAGGCAATCTCAAGAAAAGTCTCTACCTGCTCAGGGAGG	1380
1399	TACAGAGGACCTCTCAGGAATCAGACCGGTGCTGATCTCCAGGSCAGCTGTCCCTG	1458
1381	TACAGAGGACCTCTCAGGAATCAGACCGGTGCTGATCTCCAGGSCAGCTGTCCCTG	1440
1459	CTGTATCCCTGCTCCAAAGAAAGGATCTCTTAAGAGTCTCGACAGCTGAAATCTGGTT	1518
1441	CTGTATCCCTGCTCCAAAGAAAGGATCTCTTAAGAGTCTCGACAGCTGAAATCTGGTT	1500
1519	ACTACTCTCTCCAGAGCCAGCGAGTCTGGGAACTCTTAGACGCGAGTGTGTTG	1578
1501	ACTACTCTCTCCAGAGCCAGCGAGTCTGGGAACTCTTAGACGCGAGTGTGTTG	1560
1579	TGAGTGGGACCCCGTGGAGCAGAACTCTCCACAGGCTTCAGGGCTCTCTCCACCGCA	1638
1561	TGAGTGGGACCCCGTGGAGCAGAACTCTCCACAGGCTTCAGGGCTCTCTCCACCGCA	1620
1639	AGGGCAATCTCAAACTCAATGGCAAGTCTCTCCGCAAGCTTTAGAGGCACTACCCCTA	1698
1621	AGGGCAATCTCAAACTCAATGGCAAGTCTCTCCGCAAGCTTTAGAGGCACTACCCCTA	1680
1699	GCACCTTTGGCTCCCTGGACCAACTGGCTCTCTCCCATCTGAGCCCGGCGGCGCC	1758
1681	GCACCTTTGGCTCCCTGGACCAACTGGCTCTCTCCCATCTGAGCCCGGCGGCGCC	1740
1759	CCTCAGGGCTGTGAGTGGAGCAGCATCTGTCTCTCCAGTCTTTGACCAATTTGGACT	1818
1741	CCTCAGGGCTGTGAGTGGAGCAGCATCTGTCTCTCCAGTCTTTGACCAATTTGGACT	1800
1819	TGCTTGAACGCTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGGACAACCTGAGGG	1878
1801	TGCTTGAACGCTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGGACAACCTGAGGG	1860
1879	GGCTTGAACGCTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGGACAACCTGAGGG	1938
1861	GGCTTGAACGCTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGGACAACCTGAGGG	1920
1939	GTGTCTTTTCTGTGACAGCTGCCAAGAGGTGATGTCAGCCTTACAGAACCCCTAGGAA	1998
1921	GTGTCTTTTCTGTGACAGCTGCCAAGAGGTGATGTCAGCCTTACAGAACCCCTAGGAA	1980
1999	TCGTCTCAAGCTCAGCTGAGGAGAGTGGTGGCTTAGTATGGGTAGGCTCTGAGA	2058
1981	TCGTCTCAAGCTCAGCTGAGGAGAGTGGTGGCTTAGTATGGGTAGGCTCTGAGA	2040
2059	GGGTTTGCAGAGGAAACCTTGGGTGGATTCCTCCAGTGAATAGATACATCAAGGGCTCT	2118
2041	GGGTTTGCAGAGGAAACCTTGGGTGGATTCCTCCAGTGAATAGATACATCAAGGGCTCT	2100
2119	AGCTCTGACGCTGACTGAACCTGAAAGATGAGAGAAATCGCATTTGATGTGAAAGGAAT	2178
2101	AGCTCTGACGCTGACTGAACCTGAAAGATGAGAGAAATCGCATTTGATGTGAAAGGAAT	2160

2179 GGGAAACCTTCTGCTCCCGAGTGTATAGTGGGGTGGCTGAAGTGGCTACCTCTCTTGT 2238  
 |||||  
 Db GGGAAACCTTCTGCTCCCGAGTGTATAGTGGGGTGGCTGAAGTGGCTACCTCTCTTGT 2220  
 |||||  
 Qy GCGATGAGTGCACCCATGACATTTCCACCCCTGTTCTCTGGCTGACCTTCAATAGT 2298  
 |||||  
 Db GCGATGAGTGCACCCATGACATTTCCACCCCTGTTCTCTGGCTGACCTTCAATAGT 2280  
 |||||  
 Qy TTCTGTTTCCATCAACACAGAGTGTAGAACCTTGACTTCTCTGGAGGTATGTAGTG 2358  
 |||||  
 Db TTCTGTTTCCATCAACACAGAGTGTAGAACCTTGACTTCTCTGGAGGTATGTAGTG 2340  
 |||||  
 Qy ACTGCCATTTATAGAGAGAAACAGCCTCTGTGTTTCCATCTCTGCTGCTGTGATCTCA 2418  
 |||||  
 Db ACTGCCATTTATAGAGAGAAACAGCCTCTGTGTTTCCATCTCTGCTGCTGTGATCTCA 2400  
 |||||  
 Qy AAGACCTGGGAAGACTCGGACCGTGTGTGACTTCACTCAAGGGGACGATGCCCTG 2478  
 |||||  
 Db AAGACCTGGGAAGACTCGGACCGTGTGTGACTTCACTCAAGGGGACGATGCCCTG 2460  
 |||||  
 Qy GACCCCATCTTAGATCTCAGAGACTTGAACCTTGAAGCTGTCTCTAGTACCAGATGCG 2538  
 |||||  
 Db GACCCCATCTTAGATCTCAGAGACTTGAACCTTGAAGCTGTCTCTAGTACCAGATGCG 2520  
 |||||  
 Qy ATGGATGCTCTGTTTCTCAGGCCAACGGACCTAGAAATGCTGACTTATTTATTTT 2598  
 |||||  
 Db ATGGATGCTCTGTTTCTCAGGCCAACGGACCTAGAAATGCTGACTTATTTATTTT 2580  
 |||||  
 Qy GTGATCTCACTCTGTTTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 2658  
 |||||  
 Db GTGATCTCACTCTGTTTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 2640  
 |||||  
 Qy TTTTGCTGCTTCAATATGTAAGTGTGTTCTGGGAACTCCACTGTGCCACTGAAG 2718  
 |||||  
 Db TTTTGCTGCTTCAATATGTAAGTGTGTTCTGGGAACTCCACTGTGCCACTGAAG 2700  
 |||||  
 Qy TTTATGTACAGAGAAATTTGGCAATGATGTCCTCTATTCAAGGGGGTGGGGCGTT 2778  
 |||||  
 Db TTTATGTACAGAGAAATTTGGCAATGATGTCCTCTATTCAAGGGGGTGGGGCGTT 2760  
 |||||  
 Qy TTTCAATATGATGTTGAGCACTGTGGAATGAGTCTCCAGTCCCTTCAACCCCAAG 2838  
 |||||  
 Db TTTCAATATGATGTTGAGCACTGTGGAATGAGTCTCCAGTCCCTTCAACCCCAAG 2820  
 |||||  
 Qy CTGSCACCTCCCTCATCTTCTGCTGCCCCCAAAAAAAAAAAAAAAAAAAAAA 2898  
 |||||  
 Db CTGSCACCTCCCTCATCTTCTGCTGCCCCCAAAAAAAAAAAAAAAAAAAAAA 2880  
 |||||

RESULT 3  
 BC033302  
 LOCUS  
 DEFINITION Mus musculus RIKEN cDNA 1200013B22 gene, mRNA (cDNA clone MGC:30598 IMAGE:3663595), complete cds.  
 ACCESSION BC033302  
 VERSION BC033302.1 GI:23271085  
 KEYWORDS MGC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 2917)  
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullihy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Rahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blackesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalilus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 22388257  
 12477932  
 2 (bases 1 to 2917)  
 Direct Submission  
 Submitted (26-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: http://mgc.nci.nih.gov  
 Contact: MGC help desk  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
 Center code: BCM-HGSC  
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
 Contact: amg@bcm.tmc.edu  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lounseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: http://image.llnl.gov  
 Series: IRAK Plate: 40 Row: d Column: 21  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27229181.

## FEATURES

source  
 1. .2917  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="Czech II"  
 /db\_xref="taxon:10090"  
 /clone="MGC:30598 IMAGE:3663595"  
 /tissue\_type="Mammary tumor metastasized to lung."  
 MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR enhancer."  
 /clone\_lib="NCI CGAP\_Lu30"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 1. .2917  
 /genes="1200013B22Rik"  
 /note="synonym: Snark"  
 /db\_xref="LocusID:74137"  
 /db\_xref="MGI:1921387"  
 81. .1976  
 /genes="1200013B22Rik"  
 /codon\_start=1  
 /product="1200013B22Rik protein"  
 /protein\_id="AAH33302.1"  
 /db\_xref="GI:23271086"  
 /db\_xref="LocusID:74137"  
 /db\_xref="MGI:1921387"  
 gene  
 CDS  
 81. .1976  
 /translation="MESVALLQRPQASASALASARPLADGLIKSPKPLMKQAV  
 KRHHKHLNRHRYEFLTGKTYGVKKARSSGLVAIKSRKDKIKDEQDLHLIR  
 RETINSSLNHPHIAIHVEFENSISKVIMVEYASRGDLVDYISERPLSERDARHF  
 RQIVSAHLCQNGHIVRDLKLENILLDANGNIKADGLSNLYHKQKFLQTCGSPKL  
 YASPEVNGKPYGPEVDSVGLLYILVHQDHPFPDQDHLVVKQISNGA YRPPK  
 PSDACGLIRLLMNPTRATLEDVASHWVNNMGVYTGVEQBALRGHPSGDSGRA  
 SMADWLRRSSRPLLENGAKVCSFFKQHPVGGSTVPGLERQRLSKSRKENDMAQNLQ

GDPAEDTSSRPKSSLLPKGLKLLKSSSTSSGEVQEDPOELRPVDPDPGPVPAVSL  
PRKGLKSSRQRESGYSSPEPSESGELLDAVDVFSQDPVEQXSPQASGLLHRKGI  
LKLNGKSRLEGTTSTFSGSLQALASHPARPSPRGAVSEDSILSSSEFQDLDL  
PERUPETPLRGCVSDNLRLGLEPPSEGLKRWQESLGDSCFSLTDCQEVTAAYRQAL  
GICSKLS"

misc\_feature

249..1001

/gene="1200013B22Rik"

/note="S\_TKC; Region: Serine/Threonine protein kinases,  
catalytic domain. Phosphotransferases of the serine or  
threonine-specific kinase subfamily. The enzymatic  
activity of these protein kinases is controlled by  
phosphorylation of specific residues in the activation  
segment of the catalytic domain, sometimes combined with  
reversible conformational changes in the C-terminal  
autoregulatory tail"

/db\_xref="CDD:cd00180"

ORIGIN

Query Match	95.2%;	Score 2763.8;	DB 10;	Length 2917;
Best Local Similarity	98.6%;	Pred. No. 0;		
Matches 2808;	Conservative 0;	Mismatches 37;	Indels 2;	Gaps 2;

  

Qy	55	TCTGAGCCGCGGCTCAGCGCGCTGCTACTGTCGCCCGACCCCACTCCACTTCGCGGTC	114
Db	13	TCTGAGCCGCGGCTCCTCGCGCGCTGCTACTGTCGCCCGACCCCACTCCACTTCGCGGTC	72
Qy	115	CCGCAACCATGGAGTGGCGGCTTACTCCAGCGCCGCGGAGCCAGGCTCCCTCGGCTCCG	174
Db	73	CCGCAACCATGGAGTGGCGGCTTACTCCAGCGCCGCGGAGCCAGGCTCCCTCGGCTCCG	132
Qy	175	CCCTGGGCTTCGAGAGCGCCGCGGCTGGCGGAGCGGCTCATCAAGTCGCGCTAAACCTC	234
Db	133	CCCTGGGCTTCGAGAGCGCCGCGGCTGGCGGAGCGGCTCATCAAGTCGCGCTAAACCTC	192
Qy	235	TGATGAAGACGAGCGGCTGAAGCGGACCATCATCAAAACAAACCTTCGCGCACCGCTACG	294
Db	193	TGATGAAGACGAGCGGCTGAAGCGGACCATCATCAAAACAAACCTTCGCGCACCGCTACG	252
Qy	295	AGTTCTTGGAGACCTCGGCAAGGGACCTACCGGAAGGTTGMAAGGCGACGAGAGCT	354
Db	253	AGTTCTTGGAGACCTCGGCAAGGGACCTACCGGAAGGTTGMAAGGCGACGAGAGCT	312
Qy	355	CGGGCGCTCTGGTGCCATCAAGTCCATCAGGAAGACAAAATCAAAAGATGAGCAGGATC	414
Db	313	CGGGCGCTCTGGTGCCATCAAGTCCATCAGGAAGACAAAATCAAAAGATGAGCAGGATC	372
Qy	415	TGCTGCACATACGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAATCATTTG	474
Db	373	TGCTGCACATACGAGGAGATTGAGATCATGTCTTCACTCAACCAACCCCAATCATTTG	432
Qy	475	CCATCCATGAAGTGTGAGAAATAGCAGCAAGATTGTGATTGTGATGAGTATGCCAGCC	534
Db	433	CCATCCATGAAGTGTGAGAAATAGCAGCAAGATTGTGATTGTGATGAGTATGCCAGCC	492
Qy	535	GAGCGATCTGTATGATTACATCAGTGAGCGGCCACCGCTGAGTGAGCGGAGCCAGGC	594
Db	493	GAGCGATCTGTATGATTACATCAGTGAGCGGCCACCGCTGAGTGAGCGGAGCCAGGC	552
Qy	595	ATTCTTCCACAGATCGTGTGCGCTGCACTACTGCCACACAGAACGGGATCGTTCAAC	654
Db	553	ATTCTTCCACAGATCGTGTGCGCTGCACTACTGCCACACAGAACGGGATCGTTCAAC	612
Qy	655	GAGATCTCAAGCTGGAAAAATCTCTTAGATGCAATGGAACATCAAGATTGCTGACT	714
Db	613	GAGATCTCAAGCTGGAAAAATCTCTTAGATGCAATGGAACATCAAGATTGCTGACT	672
Qy	715	TTGGCTCTCCAACTGTACCAAAAGGCAAGTTCTCTCAGACGTTCTGTGGGAGCCCTC	774
Db	673	TTGGCTCTCTCAACTGTACCAAAAGGCAAGTTCTCTCAGACGTTCTGTGGGAGCCCTC	732
Qy	775	TCTACGCTCGCTGAGATGATCAACGGGAAGCCCTATGTGGGCCCGAGGTGACAGCT	834
Db	733	TCTACGCTCGCTGAGATGATCAACGGGAAGCCCTATGTGGGCCCGAGGTGACAGCT	792

Qy	835	GGTCTCTGGCGGTTCTCTCTGTATCATCTGTGTGATGCACCATGCTCCCTTTTGA	894
Db	793	GGTCTCTGGCGGTTCTCTCTGTATCATCTGTGTGATGCACCATGCTCCCTTTTGA	852
Qy	895	ATCATAAACACATGGTGAAGCAAAATCAGTAAACGGGGCTTACCGTGAGCCGCCCAAGCCGT	954
Db	853	ATCATAAACACATGGTGAAGCAAAATCAGTAAACGGGGCTTACCGTGAGCCGCCCAAGCCGT	912
Qy	955	CCGATGCTGTGGCTGATCCGGTGGCTGTTAATGTGTGAACCCCAACCCCGTCGGGCCACAC	1014
Db	913	CCGATGCTGTGGCTGATCCGGTGGCTGTTAATGTGTGAACCCCAACCCCGTCGGGCCACAC	972
Qy	1015	TGAGAGATGATAGCAGTCAATGTTGGGTCAACTCTGGGGTTACACCAACCGAGTGGGGAAC	1074
Db	973	TGAGAGATGATAGCAGTCAATGTTGGGTCAACTCTGGGGTTACACCAACCGAGTGGGGAAC	1032
Qy	1075	AGCAAGCCCTGCGTGAGGGTGGGCACCCCTAGTGTGACTTTTGGCCGGGCTCCATGGCGG	1134
Db	1033	AGCAAGCCCTGCGTGAGGGTGGGCACCCCTAGTGTGACTTTTGGCCGGGCTCCATGGCGG	1092
Qy	1135	ACTGGTTACGTCGCTCCTCGCGCCCTCTCTGGAGAAATGGAGCCAAAGGTGTGACGTTCT	1194
Db	1093	ACTGGTTACGTCGCTCCTCGCGCCCTCTCTGGAGAAATGGAGCCAAAGGTGTGACGTTCT	1152
Qy	1195	TCAAGCAGCAGCTGCCGGGAGGTGGAGACACTGTACTCTGGGCTGGAGCGGCAATTTCTC	1254
Db	1153	TCAAGCAGCAGCTGCCGGGAGGTGGAGACACTGTACTCTGGGCTGGAGCGGCAATTTCTC	1212
Qy	1255	TTAAGAAAGTCCGAAAGGAGATGACATGGCTCAAATCTGCAAGGTGACCCGGCTGAGG	1314
Db	1213	TTAAGAAAGTCCGAAAGGAGATGACATGGCTCAAATCTGCAAGGTGACCCGGCTGAGG	1272
Qy	1315	ATACCTCTTCTCGCCCTGGCAAGAGCAGCTTTAAGCTTCGAAAGGCAATTTCTCAAGAAA	1374
Db	1273	ATACCTCTTCTCGCCCTGGCAAGAGCAGCTTTAAGCTTCGAAAGGCAATTTCTCAAGAAA	1332
Qy	1375	AGTCTCTACCTCGTTCAGGGAGGTACAGAGAGACCTCAGGACCTCAGACCCGGTGCTG	1434
Db	1333	AGTCTCTACCTCGTTCAGGGAGGTACAGAGAGACCTCAGGACCTCAGACCCGGTGCTG	1392
Qy	1435	ATACTCCAGGCGAGCTGCTCCCTGCTGTATCCCTGCTCCCAAGAAAGGCAATCTTAAAGA	1494
Db	1393	ATACTCCAGGCGAGCTGCTCCCTGCTGTATCCCTGCTCCCAAGAAAGGCAATCTTAAAGA	1452
Qy	1495	AGTCTCGACAGCGTGAATCTGTTTACTCTCTCTCAGAGCCCAAGGAGTCTGGGGAAC	1554
Db	1453	AGTCTCGACAGCGTGAATCTGTTTACTCTCTCTCAGAGCCCAAGGAGTCTGGGGAAC	1512
Qy	1555	TCTTAGACGCCAGTGTGTTGTGAGTGGGGACCCCGTGGAGCAGAGTCTCCACAGG	1614
Db	1513	TCTTAGACGCCAGTGTGTTGTGAGTGGGGACCCCGTGGAGCAGAGTCTCCACAGG	1572
Qy	1615	CTTCAGGGCTCTCTCCACCGCAAGGCAATTTCTCAAACTCAATGGCAAGTCTCCCGCA	1674
Db	1573	CTTCAGGGCTCTCTCCACCGCAAGGCAATTTCTCAAACTCAATGGCAAGTCTCCCGCA	1632
Qy	1675	CAGCTTTAGAGGCACTACCCCTAGCACTTTGGGCTCTCTGGACCAATCTGCTCTCT	1734
Db	1633	CAGCTTTAGAGGCACTACCCCTAGCACTTTGGGCTCTCTGGACCAATCTGCTCTCT	1692
Qy	1735	ATCTCGACGCCCGCCAGCCGCTCAGGGGCTGTGAGTGGAGCAGACATCTCTGCTCT	1794
Db	1693	ATCTCGACGCCCGCCAGCCGCTCAGGGGCTGTGAGTGGAGCAGACATCTCTGCTCT	1752
Qy	1795	CCGAGTCTTTGACCAATTTGGACTTGGCTTGAAGCTTTCCCGAAAACCCCACTGAGGGGCT	1854
Db	1753	CCGAGTCTTTGACCAATTTGGACTTGGCTTGAAGCTTTCCCGAAAACCCCACTGAGGGGCT	1812
Qy	1855	GTGTGTCTGTGACACCTGAGGGGCTTGAAGAGCTCTCTCAGAGGTCTGAAGCGAT	1914
Db	1813	GTGTGTCTGTGACACCTGAGGGGCTTGAAGAGCTCTCTCAGAGGTCTGAAGCGAT	1872

```
QY 1915 GGTGGCAGGAATCCTTGGGGGATAGCTGCTTTTCTCTGACAGACTGCCAAGAGGTGACTG 1974
Db 1873 GGTGGCAGGAATCCTTGGGGGATAGCTGCTTTTCTCTGACAGACTGTCAAGAGGTGACTG 1932
QY 1975 CAGCCTACAGAACGCTTAGGAATCTGCTCAAGCTCAGCTCAGGAGGAGATGGTGC 2034
Db 1933 CAGCCTACAGAACGCTTAGGAATCTGCTCGAAGCTCAGCTCAGGAGGAGATGGTGC 1992
QY 2035 CTTAGTATGGGGTAGGCTCTGAGAGGTTTTCAGAGGAAACCTTGGGTGGATTCCTCCAG 2094
Db 1993 CTTAGTATGGGGTAGGCTCTGAGAGGTTTTCAGAGG-ACCTTGGGTGGATTCCTCCAG 2051
QY 2095 TGAATAGATACATCAAGGCTCTACCTGTCGAGGCTGACTGAACCTGAAAGATGAGAGA 2154
Db 2052 TGAATAGATACATCAAGGCTCTACATCTGCAAGCTGACTGAACCTGAAAGATGAGAGA 2111
QY 2155 AATCGCATTCGATCGGAAGGATGGGAACCTTGTGCGCGGAGTGTATAGTGGGGTGG 2214
Db 2112 AATCGCATTCGATCGGAAGGATGGGAGCCCTTGTGCGCGGAGTGTATAGTGGGGTGG 2171
QY 2215 CTTGAAGGTCCTACCTCTTGTGCGCATGAGTGTACCCATGACATTTTCCACCCCTGTT 2274
Db 2172 CTTGAAGGTCCTACCTCTTGTGCGCATGAGTGTACCCATGACATTTTCCCGCCCTGTT 2231
QY 2275 CTTGCGTGCACCTTCAATAGTTCCTGTTTCCATCAACACACGAGTTAGAACCTGA 2334
Db 2232 CTTGCGTGCACCTTCAATAGTTCCTGTTTCCATCAACACACGAGTTAGAACCTGA 2291
QY 2335 CTTCTGGGAGGTAATGTGTAGTACTGCTCCATTTATTAGAGAGAAACAGCCCTCTGGTTT 2394
Db 2292 CTTCTGGGAGGTAATGTGTAGTACTGCTCCATTTATTAGAGAGAAACAGCCCTCTGGTTT 2351
QY 2395 CCATCTCTGCTGTGTCATCTCAAGACCTCGGAAGACTCGGACCGCTGTTTGAATTCA 2454
Db 2352 CCATCTCTGCTGTGTCATCTCAAGACCTCGGAAGACTCGGACCGCTGTTTGAATTCA 2411
QY 2455 TCTCAAGGGGACAGATGCCCCCTGGACCCCATCTTAGATCTCAGAGACTTGAACCTTGAA 2514
Db 2412 TCTCAAGGGGACAGATGCCCCCTGGACCCCATCTTAGATCTCAGAGACTTGAACCTTGAA 2471
QY 2515 GCTGTTCTTAGTACCCAGATGTGGATGGATGCTCTGTTTCTCAGGCCAACGGACCTAGA 2574
Db 2472 GCTGTTCTTAGTACCCAGATGTGGATGGATGCTCTGTTTCTCAGGCCAACGGACCTAGA 2531
QY 2575 ATGTGCTGACTTATTTATTTTGTGATCTCACTTCTCTTTTGTGTTTGTGTTTGTGTT 2634
Db 2532 ATGTGCTGACTTATTTATTTTGTGATCTCACTTCTGTTTGTGTTTGTGTTTGTGTT 2591
QY 2635 TGTGTTGTTTGTGTTTAACTGAATT-TTGCTGCTTTTCAATAATGTGAATGCTGTGTTCT 2693
Db 2592 TGTGTTGTTTGTGTTTAACTGAATTATGCTGCTTTCAATAATGTGAATGCTGTGTTT 2651
QY 2694 GGGGAACCTCACTGCGCACTGAAGTTTATGTACAGAGAAGTATTTGGGCAATGATGCC 2753
Db 2652 GGGGAACCTCCCTCTGCGCACTGAAGTTTATGTACAGAGAAGTATTTGGCAATGATGCC 2711
QY 2754 TCTATTCAAGGGGGTGGGGGCTTTTCAAAATGTATGCTCTCAGCACTGCTCTGGATTGA 2813
Db 2712 TTTATTCAAGGGGGTGGGGGCTTTTCAAAATGTATGTTTTCAGCCCTGTTTGGATTGA 2771
QY 2814 GTCTCCAGTCCCTTCACACCAAGGCTGGCCACCTCCCTCATCTTTTCATCTGTGGCCAA 2873
Db 2772 GTTTGCACTCCCTTCACACCAAGGCTGGCCACCTCCCTCATCTTTTCATCTGTGGCCAA 2831
QY 2874 AAAAAAAAAAAAAAAAAAAAAA 2900
Db 2832 AAAAAAAAAAAAAAAAAAAAAA 2858
```

RESULT 4

AX380960

LOCUS

DEFINITION

Sequence 1 from Patent WO0212456. 2929 bp DNA linear PAT 18-MAR-2002

## ACCESSION

AX380960.1 GI:19575800

## VERSION

Rattus sp.

## SOURCE

Rattus sp.

## ORGANISM

Rattus.

## REFERENCE

1

Drucker, D.J., Rosen, C.F. and Lefebvre, D.L.

Ampk-related serine/threonine kinase, designated snrk

Patent: WO 0212456-A 1 14-FEB-2002;

1149336 ONTARIO INC. (CA)

## FEATURES

Location/Qualifiers

1..2929

/organism="Rattus sp."

/mol\_type="unassigned DNA"

/db\_xref="taxon:10118"

## ORIGIN

Query Match

Best Local Similarity

Matches 2498; Conservative

71.2%; Score 2066.8; DB 6; Length 2929;

Pred. No 0;

Mismatches 277; Indels 115; Gaps 14;

QY 49 GTGACCTCTGAGCCCGCGCTCAGCGCGCTGCTACTGCTGCCGACCCACCTCCACTC 108

Db 9 GTGACCTCTGAGCTCGCGCTCTCCGCGCTGCTGCTGCCGCGGAACTCTGGCTCCGCTC 68

QY 109 GCGGTCCCGCACCATGAGTGGTGGCTTACTCCAGGCGCCGAGCAGGCTCCCTCGG 168

Db 69 GCGGTCCCGCACCATGAGTGGTGGCTTACTCCAGGCGCGGAACTCTGGCTCCCTCGG 128

QY 169 CTTCCGCCCTGGCTCGGAGAGCGCGCGCTGCGGACCGGCTCATCAAGTCGGCTA 228

Db 129 CTTCCGCCCTGGCCACGAGAGCGCGCGCTGCGGACCGGCTCATCAAGTCGGCTA 188

QY 229 AACCTCTGATGAAGAAGCAGCGGCTGAAGCGGACCATCAAAACACAACTCGGGCACC 288

Db 189 AACCTCTGATGAAGAAGCAGCGGCTGAAGCGGACCATCAAAACACAACTCGGGCACC 248

QY 289 GCTACAGATTCTGGAGAGCTGGGCAAGGGACCTACGGGAAGGTGAAGAAGCAGCAG 348

Db 249 GCTACAGATTCTGGAGAGCTGGGCAAGGGACCTACGGGAAGGTGAAGAAGCAGCAG 308

QY 349 AGAGCTCGGGCGCTCTGTTGGCCATCAAGTCATCAGGAAGACAAAATCAAAAGATGAGC 408

Db 309 AGAGCTCGGGAGCCCTCTGTTGGCCATCAAGTCATCAGGAAGACAAAATCAAAAGATGAGC 368

QY 409 AGGATCTGTCACATACGGAGGGAGATTGAGATCATGTCTTCACTCAACACCCCCACA 468

Db 369 AGGATCTGTCACATACGGAGGGAGATTGAGATCATGTCTTCACTCAACACCCCCACA 428

QY 469 TCATTGCCATCCATGAAGTGTGAGAAATAGCAGCAAGATTGTGATGTGATGAGTATG 528

Db 429 TCATTGCCATCCATGAAGTGTGAGAAATAGCAGCAAGATTGTGATGTGATGAGTATG 488

QY 529 CCAGCGAGCGATCTGTATGATTACATCAGTCAGCGCGCCACCGCTGAGTGAGCGGACG 588

Db 489 CCAGCGAGCGATCTGTACGATTACATCAGTCAGCGCGCCACCGCTGAGTGAGCGGACG 548

QY 589 CCAGGCAATTTCTTCGACAGATCGTGTCTGCCCTGCACTACTGCCACCCAGACGGGATCG 648

Db 549 CCAGGCAATTTCTTCGACAGATCGTGTCTGCCCTGCACTACTGCCACCCAGACGGGATCG 608

QY 649 TTCACCGAGATCTCAAGCTGGAAAAATCCTTCTTAGATGCGCAATGGAACATCAAGATTG 708

Db 609 TTCACCGAGATCTCAAGCTGGAAAAATCCTTCTTAGATGCGCAATGGAACATCAAGATTG 668

QY 709 CTGACTTTGGCCCTCTCCAACTGTATCCAAAGGCAAGTTCCTCCAGACGCTTCTGTGGGA 768

Db 669 CTGACTTTGGCCCTCTCCAACTGTATCCAAAGGCAAGTTCCTCCAGACGCTTCTGTGGGA 728

QY 769 GGCCTCTCTACGCTCGCCCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCAGAGGTGG 828









Db 1801 GGGGCTGTGTCTGTGTGACAACTGAGGGGCTTGAGCAGCTCCCTCAGAGGTCTGA 1860  
Qy 1909 AGCGATGGTGGCAGGAATCCTTGGGGGATAGCTGCTTTTCTCTCAGACAGTGCAGAGG 1968  
Db 1861 AGCGATGGTGGCAGGAATCCTTGGGGGATAGCTGCTTTTCTCTCAGACAGTGCAGAGG 1920  
Qy 1969 TGACTGCAGCTCAGACAGCAAGCCCTAGGAATCTGCTCAAGCTCAGCTGAGGAGGAGGA 2028  
Db 1921 TGACTGCAGCTCAGACAGCAAGCCCTAGGAATCTGCTCAAGCTCAGCTGAGGAGGAGGA 1980  
Qy 2029 TGGTGCCCTAGTATGGGGTAGGCTCTGAGAGGGTTTGAGAGGAACC 2075  
Db 1981 TGGTGCCCTAGTATGGGGTAGGCTCTGAGAGGGTTTGAGAGGAACC 2027

RESULT 6  
AX380986  
LOCUS AX380986 2026 bp DNA linear PAT 18-MAR-2002  
DEFINITION Sequence 27 from Patent WO212456.  
ACCESSION AX380986  
VERSION AX380986.1 GI:19575826  
KEYWORDS  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1  
AUTHORS Drucker, D.J., Rosen, C.F. and Lefebvre, D.L.  
TITLE Ampk-related serine/threonine kinase, designated snark  
JOURNAL Patent: WO 0212456-A 27 14-FEB-2002;  
1149336 ONTARIO INC. (CA)  
FEATURES  
source Location/Qualifiers  
1..2026  
/organism="Rattus sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10118"

ORIGIN  
Query Match 61.0%; Score 1771.4; DB 6; Length 2026;  
Best Local Similarity 93.1%; Pred. No. 0;  
Matches 1888; Conservative 0; Mismatches 136; Indels 5; Gaps 3;  
Qy 49 GTACCTCTCAGCCGCGGCTCAGCGCGCTGCTACTGCTGCCCGACCCACTCCACTC 108  
Db 1 GTACCTCTCAGCTGCGGCTCTCCGCGGCTGCTGCTGCCCGACCCCTCCGCTC 60  
Qy 109 GCGGTCCCGCAGCATGGAGTGGTGGCTTACTTCCAGCGCCCGAGCCAGGCTCCCTCG 168  
Db 61 GCGGTCCCGCAGCATGGAGTGGTGGCTTACTTCCAGCGCCCGAGCCAGGCTCCCTCG 120  
Qy 169 CCTCCGCTGGCTCGAGAGCGCCGCGCTGCGGAGCGGGCTCATAGTGGCTA 228  
Db 121 CCTCCGCTGGCTCGAGAGCGCCGCGCTGCGGAGCGGGCTCATAGTGGCTA 180  
Qy 229 AACCTCTGATGAAGAACGAGCGGTGAAGCGGACCATCAACAAACAACCTCGGCAAC 288  
Db 181 AACCTCTGATGAAGAACGAGCGGTGAAGCGGACCATCAACAAACAACCTGAGGAC 240  
Qy 289 GCTACGAGTTCTTGAGAGCGCTGGGCAAGGCACTTACGGGAGGTGAAGAGCGACGAG 348  
Db 241 GCTACGAGTTCTTGAGAGCGCTGGGCAAGGCACTTACGGGAGGTGAAGAGCGACGAG 300  
Qy 349 AGAGCTGGGGCTGCTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGC 408  
Db 301 AGAGCTGGGGCTGCTGGCCATCAAGTCCATCAGGAAGGACAAATCAAGATGAGC 360  
Qy 409 AGGATCTGTCATACGAGGAGGAGATTGAGATCATGTTCTTCACTCAACACCCACCA 468  
Db 361 AGGATCTGTCATACGAGGAGGAGATTGAGATCATGTTCTTCACTCAACACCCACCA 420  
Qy 469 TCATTGCCATCCATGAAGTGTGTTGAGAAATAGCAAGATGTTGATTGTCATGGATG 528

Db 421 TCATTGCCATCCATGAAGTGTGTTGAGAACAGCAGCAGAGATGTTGATTGTCATGGAGTACG 480  
Qy 529 CCAGCCGAGGCGATCTGTATGATTTACATCAGTCAGCGGCCACCGCTGAGTCAGCGGACG 588  
Db 481 CCAGCCGAGGCGATCTGTACGATTAATCATCAGTCAGCGGCCACCGCTGAATGAGCGGACG 540  
Qy 589 CGAGGCATTTCTTCCAGCAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 648  
Db 541 CCAGGCATTTCTTCCAGCAGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Qy 649 TTCACCGAGATCTCAAGCTGGAACACATCTCTCTAGATGCAATGGAACATCAAGATTG 708  
Db 601 TTCACCGAGATCTCAAGCTGGAACACATCTCTCTAGATGCAATGGAACATCAAGATTG 660  
Qy 709 CTGACTTTGGCCTCTCCAACTGTACACAAAGGCAAGTTCTTCCAGACGTTCTGTGGGA 768  
Db 661 CTGACTTTGGCCTCTCCAACTGTATACAAAGGCAAGTTCTTCCAGACGTTCTGTGGGA 720  
Qy 769 GGCCTCTTACGCTGCTGCTGAGATAGTCAACGGGAGGCTTATGCTGGGCTGAGAGGTG 828  
Db 721 GGCCTCTTATGCTCTCACCTGAGATCGTCAACGGGAGGCTTATGCTGGGCTGAGAGGTG 780  
Qy 829 ACAGCTGGTCTCTGGGCGTTCTCTCTGTATCATCTCTGTGATGCAACATGCCCTTTGACG 888  
Db 781 ACAGCTGGTCTCTGGGCGTTCTCTCTGTATCATCTCTGTGATGCAACATGCCCTTTGACG 840  
Qy 889 GGAGGATCATAAACACACTGGTGAACAAATCAAGTAAACGGGCTTACCTGAGCCGCCCA 948  
Db 841 GGAGGATCATAAACACCTGGTGAACAAATCAAGTAAACGGGCTTACCTGAGCCGCTGCA 900  
Qy 949 AGCGCTCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1008  
Db 901 AACGCTCTGATGCTCTGGGCTGATCCGCTGGCTGTTTAAATGTTGAATCCCATCCGCTGG 960  
Qy 1009 CCACCTGGAGGATGTAGCAGTCACTTGGTGGTCAACTGGGCTTACACACCGGAGTGC 1068  
Db 961 CCACCTGGAGGATGTAGCAGTCACTTGGTGGTCAACTGGGCTTACACACCGGAGTGC 1020  
Qy 1069 GGGACAGGAAGCCCTCGCTGAGGGTGGGACCCCTAGTGGTGAATTTGGCGGGCTTCCA 1128  
Db 1021 GGGACAGGAAGCTCTGCGAGAGGGTGGGACCCCTAGCGGTGACTCTGGCGGGCTCTA 1080  
Qy 1129 TGGCGACTGTTTACGCTGCTCTCGCGCCCTCTGAGAGATGAGAGCAAGGCTGCA 1188  
Db 1081 TGGCGACTGTTTACGCTGCTCTCTCGCGCCCTCTCTGGAGATGGAGCAAGTGTGTA 1140  
Qy 1189 GCTTCTTCAAGCAGCACTGCTCGCGGAGGTGGAAGCACTGTACCTGGGCTGGAGCGCAAC 1248  
Db 1141 GCTTCTTCAAGCAGCACTGCTCGCGGAGGTGGAAGCACTGGAGCGGCTGGAGCGCAAC 1200  
Qy 1249 ATTCTTTAAGAGTCCCGAAGGAGATGACATGGCTCAAAATCTGAAGTGAACCCGG 1308  
Db 1201 ATTCTTTAAGAGTCCCGAAGGAGATGACATGGCTCAAAATCTGAAGTGAACCCAG 1260  
Qy 1309 CTGAGATACCTCTCTCGCCCTGGCAAGCAGACCTTAAAGCTTCCGAAAGGATTTCTCA 1368  
Db 1261 TTGAAGTACTTCTCTCGCCCTGGCAAGAACAGCTCAAGCTTCCGAAAGGATTTCTCA 1320  
Qy 1369 AGAAAAGTCTCTTACCTCGTCAAGGGAGGTACAGAGGACCTTCAAGAACTCAGACCGG 1428  
Db 1321 AGAAAAGGCTCTCTCCCTCATCGGGAGGTACAGAGGCGCTCAGGAACTCAGACCAAG 1380  
Qy 1429 TGCCTGATCTCCAGGCGAGCTGCTCCCTGCTGATCTCCCTGCTCCCAAGAAAGGATCC 1488  
Db 1381 TGTCCAATACCCAGGCGAGCTGCTCCCTGCTGATACCTCCCTGCTCCCAAGAAAGGATTC 1440  
Qy 1489 TTAAGAGTCTCGACAGCGTGAATCTGTTACTTCTCTCAGAGCCAGCGAGTCTG 1548  
Db 1441 TTAAGAGTCTCGGAGCGTGAATCTGTTACTTCTCTCAGAGCCAGCGAGTCTG 1500  
Qy 1549 GGGAACTCTTACAGCGCAGTGAATGTTTGTAGTGGGACCCCTGAGAGCAGAAAGTCTC 1608

```

Db 1501 GGGAACTCTTAGACGCGAGGTGATGTGTTGTAGTGGGAGCCCCGTGGAGCAGAGTCTC 1560
Qy 1609 CACAGGCTTACGGCTCTCTCTCCACGACGAGGCAATTTCTCAAACTCAATGGCAAGTTCT 1668
Db 1561 CACAAGCTTACGGG--CGCTCTCATGCAAGGCGATCTCTCAAACTCAATGGCAAGTTT 1617
Qy 1669 CCGCAGACGCTTAAAGGCACTACCTCTAGCACCTTTGGCTCCTCTGGACCAACTGGCCT 1728
Db 1618 CCGCAGACGCTTAAAGGCACTACCTCTAGCACCTTTGGCTCCTCTAGACCAACTGGCCT 1677
Qy 1729 CTTCCCTCTCTGACGCGCGCCGAGCCCTCAGGGGCTGTAGTGAAGACGATCC 1788
Db 1678 CCGCTCATCTTACAGCGCGGCGAGCGCTCCTCTGGAGCTGTAGTGAAGACGATCC 1737
Qy 1789 TGTCTCTCGAGTCTCTTTGACCAATTTGCACTTGGCTGCAAGCTTTCCCGAAACCCCACTGA 1848
Db 1738 TGTCTCTCGAGTCTCTTTGACCAATTTGCACTTGGCTGCAAGCTTTCCCGAAACCCCACTGA 1797
Qy 1849 GGGCTGTGTCTGTGTGGACAACTGAGGGGCTTTGAGCAGCTCTCTCTCAAGAGTCTGA 1908
Db 1798 GGAGCTGTGTCTGTGTGGACAACTGAGGGGCTTTGAGCAGCTCTCTCTCAAGAGGCTTAA 1857
Qy 1909 AGCATGCTGCGAGGATCTTTGGGGGATAGCTCTTTCTCTGACAGCTGCGCAAGG 1968
Db 1858 AACGATGCTGCGAGGATCTTTGGGGGATAGCTCTTTCTCTGACAGCTGCGCAAGG 1917
Qy 1969 TGACTGCGCTACAGCAAGCCCTAGGAATCTCTCAAGCTCAGCTGAGGAGGAGGA 2028
Db 1918 TGACAGAGCTACAGCAAGCCCTAGGAATCTCTCTGAGCTCAGCTGAGGAGGAGG 1977
Qy 2029 TGTGTCCTTAGT-ATGGGGTAGGCTCTGTAGA-GGGTTTGACAGGAAACC 2075
Db 1978 CAGTGCCCTAGTATGGGTAGACTCTTAGAGGGGTTTGACAGGAAACC 2026

RESULT 7
AC125887
LOCUS
DEFINITION
Rattus norvegicus clone CH230-274J3, *** SEQUENCING IN PROGRESS
***, 13 unordered pieces.
AC125887
AC125887.3 GI:25008671
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 291762)
Muzny,D, Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alabrooks,S., Amin,A., Auguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnshead,M., Benahmed,F.,
Blawalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesear,H., Center,A.,
Chacko,J., Chavez,D., Chen,K., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Deigado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebrgeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

```

```

Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwokedumeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parke,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Riley,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejo,Z., Usmani,K.,
Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 291762)
Worley,K.C.
Direct Submission
Submitted (02-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 291762)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23269681.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GWLD
Center clone name: CH230-274J3
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 226053 bases at least Q40
Consensus quality: 230853 bases at least Q30
Consensus quality: 233437 bases at least Q20
Estimated insert size: 233445; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is

```

\* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1
70277: contig of 70277 bp in length
70377: gap of unknown length
95031: contig of 24654 bp in length
95032: gap of unknown length
95131: gap of unknown length
95132: gap of unknown length
108956: contig of 13825 bp in length
109056: gap of unknown length
109057: contig of 162641 bp in length
171698: gap of unknown length
171797: gap of unknown length
173026: contig of 1229 bp in length
173126: gap of unknown length
174531: contig of 1405 bp in length
174631: gap of unknown length
176287: contig of 1656 bp in length
176387: gap of unknown length
177904: contig of 1517 bp in length
178004: gap of unknown length
178523: contig of 1519 bp in length
179623: gap of unknown length
180724: contig of 1101 bp in length
180824: gap of unknown length
180825: contig of 3834 bp in length
184658: gap of unknown length
184659: gap of unknown length
184758: gap of unknown length
187234: contig of 2476 bp in length
187235: gap of unknown length
187335: contig of 4428 bp in length.
Location/Qualifiers
1. .291762
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-274J3"
70378. .75596
/note="wgs_contig"
95132. .95554
/note="wgs_contig"
109057. .110574
/note="wgs_contig"

FEATURES
source
1. .291762
Query Match 54.98; Score 1593.2; DB 2; Length 291762;
Best Local Similarity 99.28; Pred. No. 0;
Matches 1601; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 55 TCTGAGCCCGCGGCTCAGCGCGGCTGCTACTGCTGCGCCGACCCACCTCCACCTCGCGGTC 114
DB 274674 TCTGAGCCCGCGGCTCTCCGGCGGCTGCTACTGCTGCGCCGACCCACCTCCACCTCGCGGTC 274733

QY 115 CCGCACCATTGAGTCCGGTGGCTTACTCAGCGCCCGAGCCAGGCTCCCTCGGCTCCG 174
DB 274734 CCGCACCATTGAGTCCGGTGGCTTACTCAGCGCCCGAGCCAGGCTCCCTCGGCTCCG 274793

QY 175 CCTGGCTCGAGAGCGCCGCGCTGCGGACGGGCTCATCAAGTCGCTAAACCTC 234
DB 274794 CCTGGCTCGAGAGCGCCGCGCTGCGGACGGGCTCATCAAGTCGCTAAACCTC 274853

QY 235 TGATGAAGAGCAGCGCGTGAAGCGGACCCATCAAAACACCAACCTCGCGGCAACGCTACG 294
DB 274854 TGATGAAGAGCAGCGCGTGAAGCGGACCCATCAAAACACCAACCTCGCGGCAACGCTACG 274913

QY 295 AGTTCTGGAGACGCTGGGCAAGGACCTACCGGAAAGGTGAAGGCAACGAGAGCT 354
DB 274914 AGTTCTGGAGACGCTGGGCAAGGACCTACCGGAAAGGTGAAGGCAACGAGAGCT 274973

QY 355 CGGGGCGCTGCTGGTGGCCATCAAGTCCATCAGGAAGACAAAATCAAAAGATGAGCAGATC 414
DB 274974 CGGGGCGCTGCTGGTGGCCATCAAGTCCATCAGGAAGACAAAATCAAAAGATGAGCAGATC 275033

QY 415 TGCTGCACATACGAGGAGGAGATTGAGATCATGTCTTCACTCAACCCACCCCATCATTTG 474

```

```

DB 275034 TGCTGCACATACGAGGAGGAGATTGAGATCATGTCTTCACTCAACCCACCATCATTTG 275093
QY 475 CCATCCATGAAGTGTGTTGAGATAGACGAAGATTGTGATTGTGATGATGATGATGATGATG 534
DB 275094 CCATCCATGAAGTGTGTTGAGATAGACGAAGATTGTGATTGTGATGATGATGATGATGATG 275153
QY 535 GAGCGGATCTGTATGATTATCATCATGAGCGGCAACGGCTGAGTGAGCGGGAACGCCAGGC 594
DB 275154 GAGCGGATCTGTATGATTATCATCATGAGCGGCAACGGCTGAGTGAGCGGGAACGCCAGGC 275213
QY 595 ATTTCTTCCACAGATCGTGTCTGCGCTGCACTACTGCCACCAAGCGGATCGTTCACC 654
DB 275214 ATTTCTTCCACAGATCGTGTCTGCGCTGCACTACTGCCACCAAGCGGATCGTTCACA 275273
QY 655 GAGATCTCAAGCTGGAACACATCTTCTAGATGCCAATGGAACATCAAGATTGCTGACT 714
DB 275274 GAGATCTCAAGCTGGAACACATCTTCTAGATGCCAATGGAACATCAAGATTGCTGACT 275333
QY 715 TTGGCCTCTCAACCTGTACCAACAAAGGCAAGTTCTTCCAGAGCGTTCTGTGGAGCCCTC 774
DB 275334 TTGGCCTCTCAACCTGTACCAACAAAGGCAAGTTCTTCCAGAGCGTTCTGTGGAGCCCTC 275393
QY 775 TCTACGCTCTCGCTGAGATAGTCAACGGGGAAGCCTATGTGGGCCCAAGGTGTGACAGCT 834
DB 275394 TCTACGCTCTCGCTGAGATAGTCAACGGGGAAGCCTATGTGGGCCCAAGGTGTGACAGCT 275453
QY 835 GGTCCTGCGGCGTTCTCTGTATCATCTGTGTGATGCGACCATGCCCTTTGACGGGCGG 894
DB 275454 GGTCCTGCGGCGTTCTCTGTATCATCTGTGTGATGCGACCATGCCCTTTGACGGGCGG 275513
QY 895 ATCATAAACACTGGTGAACCAATCAGTAAACGGGCTTACCTGAGCGCCCAAGCCGT 954
DB 275514 ATCATAAACACTGGTGAACCAATCAGTAAACGGGCTTACCTGAGCGCCCAAGCCGT 275573
QY 955 CCATGCTCTGCGCTGATCCGGTGGCTGTTAATGTGTGAACCCACCCCGCTCGGCCACAC 1014
DB 275574 CCATGCTCTGCGCTGATCCGGTGGCTGTTAATGTGTGAACCCACCCCGCTCGGCCACAC 275633
QY 1015 TGAGAGTGTAGCAGTCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1074
DB 275634 TGAGAGTGTAGCAGTCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 275693
QY 1075 AGAAGCCCTCGTGAGGTTGGGACCCCTAGTGTGCTGCTTGGCGCGGCTTCCATGGCGG 1134
DB 275694 AGAAGCCCTCGTGAGGTTGGGACCCCTAGTGTGCTGCTTGGCGCGGCTTCCATGGCGG 275753
QY 1135 ACTGGTTAGTCTGCTCTCGCGCCCTCTCTGAGAGATGAGCAAGGTGTGCAAGCTTCT 1194
DB 275754 ACTGGTTAGTCTGCTCTCGCGCCCTCTCTGAGAGATGAGCAAGGTGTGCAAGCTTCT 275813
QY 1195 TCAAGCAGCAGTGTGCGGAGGTGAAGACATGTGTACTGCGGTGAGAGCGGCAACATTCTC 1254
DB 275814 TCAAGCAGCAGTGTGCGGAGGTGAAGACATGTGTACTGCGGTGAGAGCGGCAACATTCTC 275873
QY 1255 TTAAGAGTCCCGAAGAGGAGATGATCGCTCAAAATCTGCAAGGTGACCCGCTGAGG 1314
DB 275874 TTAAGAGTCCCGAAGAGGAGATGATCGCTCAAAATCTGCAAGGTGACCCGCTGAGG 275933
QY 1315 ATACCTCTTCTCGCCCTGCAAGAGCAGGCTTAAAGCTTCCGAAGGCATTTCTCAAGAAA 1374
DB 275934 ATACCTCTTCTCGCTCTGCAAGAGCAGGCTTAAAGCTTCCGAAGGCATTTCTCAAGAAA 275993
QY 1375 AGTCTCTACCTGTCAGGGGAGGTACAGAGGACCTCAGGAACCTCAGACCGGTGCTG 1434
DB 275994 AGTCTCTACCTGTCAGGGGAGGTACAGAGGACCTCAGGAACCTCAGACCGGTGCTG 276053
QY 1435 ATACTCCAGGAGCCTGTCCTGCTGATCCTGCTCCCAAGGAAGGCATCTCTTAAGA 1494
DB 276054 ATACTCCAGGAGCCTGTCCTGCTGATCCTGCTCCCAAGGAAGGCATCTCTTAAGA 276113
QY 1495 AGTCTCGACAGCGTGAATCTGGTTACTCTCTCTCAGAGCCCAAGCAGTCTGGGGAAC 1554

```



```

QY 1702 CCTTGGCTCCCTGAGCAACTGGCTCTCTCCCATCTGAGCGCGCGCCAGCGCCCT 1761
DB 1715 CCTTGGCTCCCTGAGTGAATCGCCCACTCGCCCTCGCGCGCGCCAGCGCCCT 1774
QY 1762 CAGGGCTGTGAGTGAGGACAGCATCTCTGCTCCCGAGTCCCTTTGACCAATTTGACTTGC 1821
DB 1775 CAGGGCTGTGAGTGAGGACAGCATCTCTGCTCTGAGTCCCTTTGACCAATTTGACTTGC 1834
QY 1822 CTGAACTCTTCCCGAAACCCCACTGAGGGCTGTGTGTCTGTGGACAACTGAGGGGC 1881
DB 1835 CTGAACTCTTCCCGAAACCCCACTGAGGGCTGTGTGTCTGTGGACAACTGAGGGGC 1894
QY 1882 TTGAGAGCTCTCCCTCAGAAG-----GTCTGAAAGCATGTGTGGCAAACTCT 1929
DB 1895 TTGAGAGCTCTCCCTCAGAAGCTCTGAGGCTCTGAGGCTGTGGCGAGGATCTT 1954
QY 1930 TGGGGGATAGTCTCTTTCTCTGACAGACTGCCAAGAGGTGACTTGACGCTTACAGCAAG 1989
DB 1955 TGGGGGATAGTCTCTTTCTCTGACAGACTGCCAAGAGGTGACTTGACGCTTACAGCAAG 2014
QY 1990 CCTAGGAATCTCTCAAGCTCAGCTGAGGAGGAGATGTGCTCTAGTATGGGGTAG 2049
DB 2015 CACTGAGGGTCTGCTCAAGAGCTCACTGAGTGGAGTAGGCAATGCCCCAG-CCCGGTGAG 2073
QY 2050 GCTCTGAGAGGGTTTGAGAGGAACTCTGGTCTGAGTTCCTCC 2092
DB 2074 GCTCTGAGAGCTGTTGACCGCGAGGGAGATGCTTTC 2116

```

## RESULT 9

```

Q782778
LOCUS 3395 bp DNA linear PAT 17-MAR-2004
DEFINITION Sequence 2918 from Patent EP1396543.
ACCESSION Q782778
VERSION Q782778.1 GI:45502721
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 2918 10-MAR-2004;
RESEARCH Association for Biotechnology (JP)
FEATURES
source
1. .3395
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
118..2004
/notes="unnamed protein product".
/codon_start=1
/protein_id="CAF86008.1"
/db_xref="GI:45502722"
/translations="MSLVPARESGPTPSAELARPLASGLINKSPKPLMKQAVKRHH
HKNLHRRFEFLTKGTYGKVKARESGRLVAIKSRKOKIDQDLMIHRRRIE
IMSLNHPHILIAHEVFENSKLIVMEYASRDLYDIISERQQLSERARHFFRQIV
SAVHYQHQRVNRHVDKLENIKIADPGLSNLYHQGKFLQTFCGSPLYASP
EIVNGKPYTGPEVDSLSGLVLYLVHGTMPFDGHDHKLIVKQISNGAVREPPKPSDA
CGLIRWLLMVPNRRATLEDVASHMMWNGYATRVGEQAPHEGHPGSDSARASWAD
WLRRSRPLLENKAKVCSPEKOHAPGGSTTPCLERQHSILKKSRKENDMAQSLHSDTA
DTHARPGKSNLKLPGILKKVSAAGVQEDPPPELSIPASPGQAEPLPKGLIK
KQKRESGYSSPESESGLLDGVFVSGDPKQKPPQAGLLHRLKGLIKLNGK
SQALAEAPPTTGSDELAPPRPLARSPGAVSEDSILSESDQDLDERLPEP
PLRGCVSDNLATGLEBPPPEGSGCLRRWRQDPLGDSFCFLSDCOEVTATYQALRVC
SKLT"

```

## CDS

```

118..2004
/notes="unnamed protein product".
/codon_start=1
/protein_id="CAF86008.1"
/db_xref="GI:45502722"
/translations="MSLVPARESGPTPSAELARPLASGLINKSPKPLMKQAVKRHH
HKNLHRRFEFLTKGTYGKVKARESGRLVAIKSRKOKIDQDLMIHRRRIE
IMSLNHPHILIAHEVFENSKLIVMEYASRDLYDIISERQQLSERARHFFRQIV
SAVHYQHQRVNRHVDKLENIKIADPGLSNLYHQGKFLQTFCGSPLYASP
EIVNGKPYTGPEVDSLSGLVLYLVHGTMPFDGHDHKLIVKQISNGAVREPPKPSDA
CGLIRWLLMVPNRRATLEDVASHMMWNGYATRVGEQAPHEGHPGSDSARASWAD
WLRRSRPLLENKAKVCSPEKOHAPGGSTTPCLERQHSILKKSRKENDMAQSLHSDTA
DTHARPGKSNLKLPGILKKVSAAGVQEDPPPELSIPASPGQAEPLPKGLIK
KQKRESGYSSPESESGLLDGVFVSGDPKQKPPQAGLLHRLKGLIKLNGK
SQALAEAPPTTGSDELAPPRPLARSPGAVSEDSILSESDQDLDERLPEP
PLRGCVSDNLATGLEBPPPEGSGCLRRWRQDPLGDSFCFLSDCOEVTATYQALRVC
SKLT"

```

## ORIGIN

```

Query Match 47.4%; Score 1375.8; DB 6; Length 3395;
Best Local Similarity 81.7%; Pred. No. 0;

```

```

Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;
QY 82 CTACTGCTGCGGACCCCACTTCAACCTCGCGGTCGCCGACCATGAGTGGTGGCTTAC 141
DB 77 CTACTGATTTCCCTGCGGCCCTTGCTCACTCTGCTCGCCATGAGTGCCTGTTTCG 136
QY 142 TCAGCGCGCGGAGCAGGCTCCCTCGGCTTCGCGCTCGGCTCGGAGAGCGCCGCGCGC 201
DB 137 CGCGGCGCTCCGCGCCCACTTCCCTCGGCGCGAGAGCTA-----GCCGCGCGC 184
QY 202 TGCGGACCGGCTCATCAAGTCGCTAAACCTCTGATGAAGAGCAGGCGGTGAAGCGGC 261
DB 185 TGCGGAGGCGGTGATCAAGTCGCGCAGGCCCTTAATGAAGAGCAGGCGGTGAAGCGGC 244
QY 262 ACCATCAAAACAAACCTTGCAGCACCCTGACGATCTCTGGAGACGCTGGGCAAGGCA 321
DB 245 ACCACCAAGACAACTCGCGCACCTGAGATCTCTGGAGACCTTGGGCAAGGCA 304
QY 322 CTTACCGGAGGTGAAGAGGACGACGAGTCTCGGCGCTCTGTTGGCCATCAAGTCCA 391
DB 305 CTTACCGGAGGTGAAGAGGCGCGGAGCTCGGCGCGCTTGTGGCCATCAAGTCAA 364
QY 382 TCAGGAAAGACAAATCAAGATGAGCAGGATCTGTCACATACGAGAGGAGATTGAGA 441
DB 365 TCCGGAAGGACAAATCAAGATGAGCAGGATCTGATGACATACGAGGAGATTGAGA 424
QY 442 TATGTTCTTCACTAAACCAACCCCACTCATTTGCGCATCCATGAAGTGTGAGAAATGAGA 501
DB 425 TCATGTCTCATCACTAACCAACCCCTCACATCAATTTGCCATCCATGAAGTGTGAGAAATGAGA 484
QY 502 GCAAGATTCTGATGTTCATGGAGTATGCCAGCGGAGGCGATCTGTATGATTATCATCAGTG 561
DB 485 GCAAGATCTGATGTCATGAGTATGCGAGCGCGGAGCTCGGCGCGCTTGTGGCCATCAAGTCAA 544
QY 562 AGCGGCCACGCTGAGTGAGCGGAGCGCGAGGCAATTTCTTCCGACAGATCGTGTCTGCGCC 621
DB 545 AGCGGCAGCAGCTCAGTGAGCGGAGCTAGGCAATTTCTTCCGCGCAGATCGTCTCTGCGCG 604
QY 622 TGCACTACTGCCACCAAGAGCGGATGTTTACCGGAGATCTCAAGCTGGGAAACATCTCTTC 681
DB 605 TGCACTATTGGCCATCAGAAACAGAGTTGTCCACCGAGATCTCAAGCTGGGAAACATCTCTCT 664
QY 682 TAGATGCCAATGGAACATCAAGATTGCTGATCTTTGGCTCTTCCAACTGTACCAACAAAG 741
DB 665 TGGATGCCAATGGAATATCAAGATTGCTGATCTTGGCTCTTCCAACTGTACCATCAAG 724
QY 742 GCAAGTCTCTCAGACGTTCTGTGGAGCCCTCTCTAGCGCTCGCTGAGATAGTCAACG 801
DB 725 GCAAGTCTCTGACAGACATTTCTGTGGAGCGCCCTCTATGCTTCCGCGCAGAGATTGCAATG 784
QY 802 GGAAGCCCTATGTGGGCGCAGAGGTGGACAGCTGCTCTGCGCGCTTCTCTCTGATACATCC 861
DB 785 GGAAGCCCTTACACAGCGCCAGAGGTGGACAGCTGCTGCTTCTCTCTCTCATATCC 844
QY 862 TGGTGCATGGCAACCATGCTTTTGAAGCGGAGGATCATAAAACTGGTGAAGCAAAATCA 921
DB 845 TGGTGCATGGCAACCATGCTTTTGAAGCGGATGACATTAAGATCTTAGTGAAGCAAGATCA 904
QY 922 GTAAAGCGGCTTACCGTGAGCGCGCCCAAGCGCTCGATGCTGTGCTGTGATCCCGTGGC 981
DB 905 GCAAGCGGCGCTTACCGGAGGACCACTAAACCTCTGATGCTGTGCTGTGATCCCGTGGC 964
QY 982 TGTAAATGTTGAACCCCACTCGGCGGACATCGAGGATGATGAGGATGATGAGGATGATGAGG 1041
DB 965 TGTAAATGTTGAACCCCACTCGGCGGACATCGAGGATGATGAGGATGATGAGGATGATGAGG 1024
QY 1042 TCAACTGGGGTTTACCAACCGGAGTGGGGAAACAGAGAGACCTTGGCTGAGGCTGGGCGACC 1101
DB 1025 TCAACTGGGGTTTACCGGAGTGGGGAAACAGAGAGACCTTGGCTGAGGCTGGGCGACC 1084
QY 1102 CTAGTGGTGAATTTGGCGGGGCTTCATGGCGGACTGGTGTAGCTGCTCTCGCGCCCC 1161
DB 1085 CTGGCAGTGAATCTGCGCGCGCTCCATGGGTGACTGGCTCGCGGCTTCTCTCTCTCTCTCTCT 1144

```



```
Qy 1162 TCCTGGAGAAATGGAGCCCAAGGTGTGCAGCTTCTTCAAGCAGCAGCGTCCGGAGGTTGAA 1221
Db 1145 TCCTGGAGAAATGGGGCCCAAGGTGTGCAGCTTCTTCAAGCAGCAGCATGCACCTGGTGGGGAA 1204
Qy 1222 GCATGTACCTGGCTGGAGGGGCAATCTCTTTAAGAGTCCCGAAGAGAGATGACA 1281
Db 1205 GCACACCCCTGGCTGGAGCGCCAGCATTCGTCTCAAGAAGTCCCGAAGAGAGATGACA 1264
Qy 1282 TGGCTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCTCTCTCGCCCTGGCAAGAGCA 1341
Db 1265 TGGCCAGTCTCTCCACAGTGACAGCGTGATGACATCGCCATCGCCCTGGCAAGAGCA 1324
Qy 1342 GCCTTAAGCTTCCGAAGGCAATCTTCAAGAAAGTCTCTTACCTCGTCRAGGGAGGTAC 1401
Db 1325 ACCTCAAGCTGCCAAAGGCAATCTCAAGAAGAGGTGTGAGCCTCTGCGAAGGGGTAC 1384
Qy 1402 AGGAGGACCTCAGGAATCAGACCGGTGCTGATCTCTCAGGGCAGCCTGTCTCCCTGTG 1461
Db 1385 AGGAGGACCTCAGGAGCTCAGGCCAATCCCTGCGAGCCAGGGCAGGCTGCC----- 1438
Qy 1462 TATCCCTGCTCCCAAGGAAAGGCAATCTTAAGAGTCTCGACAGCGTGAATCTGGTTACT 1521
Db 1439 ---CCCTGCTCCCAAGAGGGCAATCTCAAGAAGCCCGACAGCGGAGTCTGGCTACT 1495
Qy 1522 ACTCTCTCCAGAGCCCGAGGAGTCTGGGAACTCTTAGAGCCGAGTGATGTGTTGTGA 1581
Db 1496 ACTCTCTCCGAGCCCGAGTCTGGGAGCTCTGGAGCGAGCGACGTGTTGTGA 1555
Qy 1582 GTGGGAGCCCGTGGAGCAGAGTCTCCAGGCTTCAGGGCTCCTCTCAGCGCAAGG 1641
Db 1556 GTGGGAGTCCCAAGGAGCAGAGCTCCGGAAGCTTCAGGGCTGCTCCTCATCGCAAG 1615
Qy 1642 GCATCTCAAACTCAATGGCAAGTCTCCGCAAGCTTAGAGGCACTACCCCTAGCA 1701
Db 1616 GCATCTCAAACTCAATGGCAAGTCTCCGAGAGCTCTGGAGCTCGCGGCCCCACCA 1675
Qy 1702 CCTTTGGCTCCCTGGACCAACTGGCTCTCTCCATCTCGAGCCCGCCAGCGCCCT 1761
Db 1676 CCTTGGCTCCCTGGATGAATCGCCCCACCTCGCCCCCTGGCCCCGGCCAGCGACCT 1735
Qy 1762 CAGGGCTGTGAGTGAGCAGCATCTGTCTCCGAGTCTCTTGGACCAATGAGCTTGC 1821
Db 1736 CAGGGCTGTGAGTGAGCAGCAGCATCTGTCTCTGAGTCTCTTGGACCAATGAGCTTGC 1795
Qy 1822 CTGAACGTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGGAACAACCTGAGGGGC 1881
Db 1796 CTGAACGGCTCCAGAGCCCCCACTGCGGGGCTGTGTCTGTGGAACAACCTCAGGGGC 1855
Qy 1882 TTGAGCAGCTCTCCCTCAGAAAG-----GTCTGAAGCGATGGTGCGAGGAATCCT 1929
Db 1856 TTGAGGAGCCCCCTCAGAGGGCCCTTGAAGCTGCTTGAGGCGTGGCGGAGATCCTT 1915
Qy 1930 TGGGGATAGTGCTTTTCTTGACAGACTGCCAAGAGGTGACTGACGCTTACAGACAAG 1989
Db 1916 TGGGGGACAGTGCTTTTCTTGACAGACTGCCAAGAGGTGACAGCAGCTTACCGACAGG 1975
Qy 1990 CCTAGGAATCTGCTCAAGCTCAGCTGAGAGGGAGATGCTGCCCTAGTATGGGGTAG 2049
Db 1976 CACTGAGGGTGTGCTCAAGCTCAGCTGAGTGAGTAGGATGAGGATGAGGCTTCC 2034
Qy 2050 GCTCTGAGAGGGTTTGAGAGGAACCTCGGTGCGATTCTCTCC 2092
Db 2035 GCTCTGAGTGAGCTGTTGACCCCGAGGGGAGATGCTTCTC 2077
```

RESULT 10  
LOCUS BD127309 3395 bp DNA linear PAT 18-SEP-2002  
DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
ACCESSION BD127309  
VERSION BD127309.1 GI:23222254  
KEYWORDS JP 2002017375-A/2740.

```
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3395)
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002017375-A/2740
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
PI SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof PH Key
Location/Qualifiers
(118)..(2001).
FT CDS
source Location/Qualifiers
1..3395
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 47.4%; Score 1375.8; DB 6; Length 3395;
Best Local Similarity 81.7%; Pred. No. 0;
Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;
Qy 82 CTACTGTGCGCGGACCCACCTCCACCTCGGGTCCGCGACCATGGAGTCCGTGGCCCTTAC 141
Db 77 CTACTGATTCCTCGCGCCCTTGTCTCACTCTGCTCGCCATGGAGTCCGTGTTTCG 136
Qy 142 TCAGAGCCCGAGCCAGGCTCCCTCGGCTTCGCGCTTCGAGAGCGCCGCGCGC 201
Db 137 CGCGGCTCCGCGCCCTCCCTCGCGCGCAGAGCTA-----GCCCGCGCGC 184
Qy 202 TGGCGAGCGGCTCATCAAGTCCCTAACTCTGATGAAGAAGCAGCGCGTGAAGCGC 261
Db 185 TGGCGAAGGGCTGATCAAGTCCGCCCAAGCCCCCTAATGAAGAAGCAGCGCGTGAAGCGC 244
Qy 262 ACCATCAACAACAACCTCGCGCACCGCTACGAGTTCCTGGAGAGCGTCGGCAAGGCA 321
Db 245 ACCACCAAGCAACACCTCGCGCACCGCTACGAGTTCCTGGAGAGCGCTGGCAAGGCA 304
Qy 322 CCTACGGGAAGGTGAAGAAGGCAAGAGAGTCTCGGGCGCTCTGGTGGCCATCAAGTCCA 381
Db 305 CCTACGGGAAGGTGAAGAAGGCGCGGAGAGTCTCGGGCGCTCTGGTGGCCATCAAGTCAA 364
Qy 382 TCAGGAAGCAAAATCAAAAGATGAGCAGGATCTGTGCAATACGAGGAGATTGAGA 441
Db 365 TCCGAAGGCAAAATCAAAAGATGAGCAGGATCTGTGCAATACGAGGAGATTGAGA 424
Qy 442 TCATGCTTTCACCTCAACCCACCTCATCTCCCATCCATGAAGTGTGTTGAGAATAGCA 501
Db 425 TCATGCTATCCTCAACCCCTCATCTCCCATCCATGAAGTGTGTTGAGAATAGCA 484
Qy 502 GCAAGATTGTGATTGTTCATGGAGTATGCCAGCGGAGCGGATCTGTATGATTACATCAGTG 561
Db 485 GCAAGATCTGTATCGTTCATGGAGTATGCCAGCGGCGGACCTTTATGACTACATCAGCG 544
Qy 562 AGCGGCAAGCTGTGAGTGAGGAGCGCGCAGGATTTCTTCCGACAGATCGTCTGCC 621
Db 545 AGCGGAGCAGCTCAGTGAGCGCGAAGCTAGGCAATTTCTTCCGCGAGATCGTCTGCCG 604
```



QY	622	TGCACTACTGCGCCACAGAA	CGGATCGTTTCA	CCGAGATCTCA	AGCTGGA	AAACATCTTC	681			
Db	605	TGCACTATTGCAATCAGAA	CAGATTGTTC	CCACCGAGATCT	CAAGCTGG	AGAACATCTCTCT	664			
QY	682	TAGATGCCAATGGAAACAT	CAAGATTGTG	ACTTTGG	CCCTCTC	CAACACCTGTACCAACAAG	741			
Db	665	TGGATGCCAAATGGGAATAT	CAAGATTGTG	ACTTCGG	CCCTCTC	CAACACCTCTTACCATCAAG	724			
QY	742	GCAAGTTCCCTCAGACGTT	CTGTGGG	AGCCCTCTCT	ACGCCTCG	CTGAGATAGTCAACG	801			
Db	725	GCAAGTTCCCTCAGACAT	TTCTGTGGG	AGCCCTCTAT	GCTCCTGCC	CAGAGATTTGTCAATG	784			
QY	802	GGAAGCCCTATGTGGGCC	CACAGAGTGG	CAGCTG	GTCTCTGGG	GGTTCTCTCTGTACATCC	861			
Db	785	GGAAGCCCTACACAGGCC	CACAGAGTGG	CAGCTG	GTCTCTGGG	GGTTCTCTCTATCATCC	844			
QY	862	TGGTGCATGGCACCATG	CCCTTTTGA	CGGGCAGGATCA	TAA	AACA	CTGGTGAAGCAAAATCA	921		
Db	845	TGGTGCATGGCACCATG	CCCTTTTGA	TGGGCATG	ACCA	TAA	AGATCTTAGTGAAA	CAGATCA	904	
QY	922	GTAACGGGGCTTACCG	TGAGCCGCC	AAGCCGTCG	ATGCTGTGG	CCCTGATCCGGTGGC	981			
Db	905	GCAACGGGGCTACCG	GAGCACTAA	ACCCTCTG	ATGCTGTGG	CCCTGATCCGGTGGC	964			
QY	982	TGTTAAATGGTGAAC	CCCCACCCGTCGG	GCAC	CTGGAGATGTAG	CCAGTCATTTGGTGGG	1041			
Db	965	TGTTGATGGTGAAC	CCCCACCCGTCGG	GCAC	CTGGAGATGTGG	CCAGTCAC	TGGTGGG	1024		
QY	1042	TCAACTGGGGTTAC	CACACCGGAGT	CGGGAA	CAGGAAGCCCTCG	GTGAGGTGGG	CACC	1101		
Db	1025	TCAACTGGGGCTAG	CCGCA	CCCGATGGG	AGACGAGGCTCC	GCATGAGGGTGGG	CACC	1084		
QY	1102	CTAGTGTGTGACTTT	TGGCGGGCTCC	ATGGCGG	ACTGGT	TACGTGCGTCTCG	CGCCCCC	1161		
Db	1085	CTGCGAGTGACTCT	GCCCGCGCTCC	ATGGCTG	ACTGGCTCGG	GGTTCTCTCC	CGCCCCC	1144		
QY	1162	TCCTGGAGAAATGG	AGCCAAAGGTGTG	CAGCTTTCTT	CAAGCAGC	ACGTGCCCGGAGGTGGAA	1221			
Db	1145	TCCTGGAGAAATGGG	GCCAAAGGTGTG	CAGCTTTCTT	CAAGCAGC	ATGCACTTGTGGGGGAA	1204			
QY	1222	GCATGTACCTGGGCT	GGAGCGGCAAC	ATCTCTT	TAAGAAGTCCG	CAAGAGGAATGACA	1281			
Db	1205	GCACACCCCTGG	CCTGGAGCGC	CAGATTCG	CTCAAGATCCG	CAGAGGAATGACA	1264			
QY	1282	TGGCTCAAAATCTG	CAAGGTGAC	CCCGCTGAG	GAATACCTTT	CTCGCCCTGGC	CAAGACA	1341		
Db	1265	TGGCCCACTCTCT	CCACAGTGC	ACGCTGAT	GACATG	CCCATCGCCCTGG	CAAGACA	1324		
QY	1342	GCCTTAAGCTTCC	GAAGGCAATCT	CAAGAA	AAAGTCCTT	ACTCTGT	CAGGGAGGTAC	1401		
Db	1325	ACCTCAAGCTG	CCAAAGGCAATCT	CAAGAA	AAAGGTGT	CAGCCTCTG	CAAGAGGGGTAC	1384		
QY	1402	AGGAGGACCTCT	CAGGAAC	CTCAG	ACGGGTGC	CTGATATCC	AGGGCAGCCTGTCC	CTGCTG	1461	
Db	1385	AGGAGGACCTCT	CGGAGCTCAG	CCCAATCCCT	CTGGAGCC	CCAGGCGGTGCC	-----	1438		
QY	1462	TATCCCTGTCTCC	AAAGAAAGGCAAT	CCCTTA	AGAA	GTCTG	CA	ACGGTGAATCTGGTTACT	1521	
Db	1439	---CCCTGTCTCC	AAAGAGGCAATCT	CTCA	AGAA	AGCCCG	CAC	AGCGAGTGTGTTGTGA	1495	
QY	1522	ACTCTCTCTCAG	AGCCAGCAGTCT	CTGGG	AACTCTT	TA	AG	AGCGAGTGTGTTGTGA	1581	
Db	1496	ACTCTCTCTCC	AGCCAGCAGTGAAT	CTGGG	AGCTTT	TG	AG	CAGGCGAGTGTGTTGTGA	1555	
QY	1582	GTGGGACCCGCTGG	AGCAGAAGTCT	CC	ACAGGCTT	CAGGGCTCTCTCT	CC	ACCGCAAGG	1641	
Db	1556	GTGGGATCC	AAAGGAGCAGAAG	CCCTCG	CAAGCTT	CAGGGTGTCT	CCTCC	ATCGCAAG	1615	
QY	1642	GCATTC	CAAACTCAATGG	CAAGTTCT	CCG	CA	CAGCCTTAG	AGGCAC	TACCCCTAGCA	1701
Db	1616	GCATTC	CAAACTCAATGG	CAAGTTCT	CC	CA	CAGCCTTAG	AGGCAC	TACCCCTAGCA	1675

Qy	1702	CCCTTGGCTCCCTGGACCAACTGGCTCTCCATCCTGTGAGCCGGCCGACGGCCCTT	1761
Db	1676	CCCTTGGCTCCCTGGATGAACCTGGCCCACTCGCCCCCTGGCCCGGGCCGACGGACCTT	1735
Qy	1762	CAGGGGCTGTGAGTGAGGACAGCATCTGTTCCTCCAGTCTCTTTGACCAATTTGACTTGC	1821
Db	1736	CAGGGGCTGTGAGGAGGACAGCATCTGTTCCTCTGAGTCTCTTTGACCACTGACTTGC	1795
Qy	1822	CTGAACGTCTTCCGGAACCCCACTGAGGGGCTGTGTCTGTGTGGAACAACCTGAGGGGGC	1881
Db	1796	CTGAACGGCTCCGAGAGCCCCCACTGGGGGCTGTGTGTCTGTGTGGAACAACCTCAGCGGGC	1855
Qy	1882	TTGAGCAGCTCCTCTCAGAAAG-----GTCTGAAGCGCATGTGTGCGCAGGAATCCT	1929
Db	1956	TTGAGGAGCCCCCTCAGAGGGCCCTTGAAGCTGCCTGAGGCGCTGGCGCAGATCCTT	1915
Qy	1930	TGCGGGATAGCTGCTTTTCTCAGACAGTGCACAAGAGTGACTGACGCTCAGACACAAG	1989
Db	1916	TGCGGGACAGCTGCTTTTCCCTGACAGACTGCCAGGAGGTGACAGCGACTTACCGACAGG	1975
Qy	1990	CCCTAGGAATCTCTCAAGCTCAGCTGAGGAGGAGATGGTGCCTTATGCGGTAG	2049
Db	1976	CCTGAGGGTCTCTCAAGCTCAGCTGAGTGGAGTAGGCAITGCCCCAG-CCCGGTGAG	2034
Qy	2050	GCCTCTGAGAGGGTTTGACAGGAAACCTTGGTTCGGATTCTCTCC	2092
Db	2035	GCCTCTGAGTGCAGTGGTTTGACCCCGAGGGGAGATGCCTTC	2077
RESULT 11			
AK074830			
LOCUS	AK074830	3395 bp mRNA linear	PRI 03-SEP-2002
DEFINITION	Homo sapiens cDNA FLJ90349 fis, clone NT2RP2003179, weakly similar to PUTATIVE SERINE/THREONINE-PROTEIN KINASE EMK (EC 2.7.1.-).		
ACCESSION	AK074830		
VERSION	AK074830.1	GI:22760531	
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T., Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S., Kawai-Hito, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Kojima, S., Nagahara, K., Masuho, Y., Ono, T., Okano, K., Yoshikawa, Y., Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and Ninomiya, K.		
TITLE	NEDO human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3395)		
AUTHORS	Isogai, T. and Otsuki, T.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan		
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).		
FEATURES			
source	Location/Qualifiers		
	1..3395		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="NT2RP2003179"		
	/cell_line="NT2"		
	/cell_type="teratocarcinoma"		
	/clone_lib="NT2RP2"		
	/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal		

precursor cells after 2-weeks retinoic acid (RA)

induction"

118..2004

CDS

/notes="unnamed protein product"

/codon\_start=1

/protein\_id="BAC11234.1"

/db\_xref="GI:22760532"

translation="MESLVFARRSGPTPSAAELARPLAEGLIKSPKPMKQAVKRRHH  
HKNLRHYEFLTLGKTYGVKVKARESSGRIVAKSIRKDKIKDQDLMHIREITE  
IMSLNHPHIIAIEHPENSSKIVIMEYASRGDLXYISERQOLSREARHPFRQIV  
SAVYCHONRVHRDLKENILLDANGNIKIADPEGLSNLYHQKFLQTCGSPLYASP  
EIVNGKPYTPGVDSNLSGLVLLYLVHGTMPFDGHDHKLIVKQISNGAYREPXPDSDA  
CGLIRLLWMTNRATLEADVASHWVWNGYATRVGQEAHPGHEGHPGSDSARASNAD  
WLRSSRPLENGAKVCSFFKQAPAGGGSTTGLERQHSLSKSKIKENDMAQSLHSDTA  
DDTAHRPKNLKLPKGLIKKVSASAEQVEDPPELSPASPQOAPLLPKKGLIK  
KPRORESGYSSPEPSSGELLADGVFVSDPKQKPPQOASGLLHRLKGLIKLNGKF  
SOTALELAAPTTPGSLDELAPPRLARASPSGAVSEDSILSSEFQDLPLPERLPEP  
PLGCVSDNLTGLUEPPEPSEPGSCLRRKQDPUGDSCFSLTDCQEVATYRQALRVC  
SKLT"

ORIGIN

Query Match 47.4%; Score 1375.8; DB 9; Length 3395;

Best Local Similarity 81.7%; Pred. No. 0;

Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;

Qy	82	CTACTGTCGCGGACCACTCCACCTCGCGGTCCCGCACCATGAGTTCGGTGCCTTAC	141
Db	77	CTACTGATTCCTCGCGCCCTTCTCACTCTCTGCTCGCATGAGTTCGCTGTTTCG	136
Qy	142	TCAGGCGCGGACGAGGTCCTCGGCTCGGCTCGGCTCGGCTCGGAGAGCGCGCGGC	201
Db	137	CGGCGGCTCGGCGCCCACTCTCTCGGCGCGCAGAGCTA-----GCCCGCGCGC	184
Qy	202	TGCGGACGCGGCTCATCAAGTCGCTAAACCTCTGTATGAAGACGACGCGGTGAAGCGC	261
Db	185	TGCGGAGGCGGTGATCAAGTCGCGCCCAAGCCCTTAATGAAGACGACGCGGTGAAGCGC	244
Qy	262	ACCATCAAAACACAACTCGGCGCACCGCTACGAGTTCTTGGAGACGCTCGGCAAGGCGA	321
Db	245	ACCACCAACAGCAACCTCGGCGCACCGCTACGAGTTCTTGGAGACGCTCGGCAAGGCGA	304
Qy	322	CCTACGCGGAAGGTGAAGAGGCGACGAGAGCTCGGCGGCTCTGTTGGCCATCAAGTCCA	381
Db	305	CCTACGCGGAAGGTGAAGAGGCGCGGAGAGCTCGGCGGCTCTGTTGGCCATCAAGTCCA	364
Qy	382	TCAGGAAGACAAATCAAGATGACGAGGATCTGTCGACATACGAGGAGGAGATTGAGA	441
Db	365	TCGGAAGGACAAATCAAGATGACGAGGATCTGATGACATACGAGGAGGAGATTGAGA	424
Qy	442	TCATGTCCTCACTCAACACCCACATCATTTGCCATCCATGAAGTGTGTTGAGAAATAGCA	501
Db	425	TCATGTCATCACTCAACACCCCTCACATCATTTGCCATCCATGAAGTGTGTTGAGAAATAGCA	484
Qy	502	GCAAGATTGATTGTCATGAGATATGCCAGCGGAGCGATCTGTATGATTATCATCAGTG	561
Db	485	GCAAGATCGTGATCGTCATGAGATATGCCAGCGGCGGAGCTTTATGACTACATCAGCG	544
Qy	562	AGCGGCCACGCGTCAGTGAGCGGAGCGCAGGCAATTTCTTCGACAGATCGTCTGCCC	621
Db	545	AGCGGCAGAGCTCAGTGAGCGGAGCTAGGCAATTTCTTCGCGCAGATCGTCTGCGCG	604
Qy	622	TGCACTACTGCCACCAAGCGGATCGTTCAACGAGATCTCAAGCTGGAAGAAATCCTCTC	681
Db	605	TGCACTATTGCCATCAGAACAGAGTGTCCACCGAGATCTCAAGCTGGAAGAAATCCTCTC	664
Qy	682	TAGATGCAATGGAACATCAAGATTGCTGACTTTGGCTCTCCAACTGTACCAAAAG	741
Db	665	TGGATGCAATGGAATATCAAGATTGCTGACTTTGGCTCTCCAACTGTACCAATCAAG	724
Qy	742	GCAAGTTCCTCCAGACGTTCTGTGGAGCCCTCTCTACGCTCGCCTGAGATAGTCAAG	801
Db	725	GCAAGTTCCTCCAGACATCTGTGGAGCCCTCTCTATGCTCGCCAGAGATTGTCAATG	784

Qy	802	GGAGAGCCTATGTGGCCGAGAGTGGACAGCTGGTCTCTGGGGCTTCTCCTGTACATCC	861
Db	785	GGAGCCCTACACAGGCCGAGAGTGGACAGCTGGTCCCTGGGTGTTCTCCTCTACATCC	844
Qy	862	TGTCATGTCACCATGCTCTTTGAACGCGGAGGANTATAAAACACTGGTGAAGCAATCA	921
Db	845	TGTCATGTCACCATGCTCTTTGATGGGATGACCATTAAGATCTCTAGTGAAGACAGATCA	904
Qy	922	GTAACGGGGCTTACCGTGAGCGCCCAAGCCGTCGATGCTGTGGCTCATCCGGTGGC	981
Db	905	GCAACGGGGCTTACCGGGAGCCACTTAACCTCTGATGCTGTGGCTCATCCGGTGGC	964
Qy	982	TGTTAATGTGAACCCCAACCCGTCGGGCCACACTGGAGGATGTAGCCAGTCAATTTGGTGG	1041
Db	965	TGTTGATGTGAACCCCAACCCGTCGGGCCACACTGGAGGATGTGGCCAGTCACTGGTGG	1024
Qy	1042	TCACTGGGGTTACACACCGGAGTCGGGGAACAGGAAGCCCTGCTGTAGGCTGGGCAAC	1101
Db	1025	TCAACTGGGGCTACGCCACCCGAGTGGGAGAGGAGGCTCCGCATGAGGGTGGGCAAC	1084
Qy	1102	CTAGTGTGACTTTGGCCGCGCTTCATGCGGACTGGTTACGTCGCTCGCGCCCCC	1161
Db	1085	CTGCGAGTGACTTTCGCCGCGCTTCATGCTGCTCGGCTCGCGCTTCCTCCGCCCCC	1144
Qy	1162	TCCTGGAGAAATGGAGCAAGTGTGCAGCTTCTTCAAGCAGCAGCTGCGCGGAGGTGGA	1221
Db	1145	TCCTGGAGAAATGGAGCAAGTGTGCAGCTTCTTCAAGCAGCAGCTGCGCGGAGGTGGA	1204
Qy	1222	GCACTGTACTGGGCTGGAGCGGCAACATTTCTCTTAAGAAAGTCCCGAAGAGGATGACA	1281
Db	1205	GCAACACCTTGGCTGGAGCGCGCAGCTTCTTCAAGAAAGTCCCGAAGAGGATGACA	1264
Qy	1282	TGCTCAAAATCTGCAAGGTGACCGCTGAGATACCTCTTCTCGCCCTGGCAAGAGCA	1341
Db	1265	TGCGCCAGTCTCTCCAGTGACACGCTGATGACATGCGCCCTGGCAAGAGCA	1324
Qy	1342	GCCTTAAGCTTCCGAAAGGCAATTTCTCAAGAAAAAGTCTCTACCTCTGTCAGGGGAGTAC	1401
Db	1325	ACCTCAAGCTGCCAAAGGCAATTTCTCAAGAAAGGCTGTCAGCTCTGCAAGAGGGTAC	1384
Qy	1402	AGAGAGCCTCTAGGAACCTCAGACCGGCTGCTGATCTCAGGGGAGCCTGTCTCTGCTG	1461
Db	1385	AGAGAGCCTCTCGGAGCTCAGCCCAATCTCTCGAGCCCGAGGCGAGCTGCTGCTGCTG	1438
Qy	1462	TATCCCTGCTCCCAAGGAAGGCAATCTTAAAGAGTCTGCACAGCTGATCTGTTACT	1521
Db	1439	---CCCTGCTCCCAAGGAAGGCAATTTCTCAAGAAAGCCCGACAGCGGAGTCTGCTGCT	1495
Qy	1522	ACTCTCTCCAGAGCCGAGCTCTGGGAACTCTTAGACGCGAGTGTGTTGTGA	1581
Db	1496	ACTCTCTCCGAGCCGAGTGAATCTGGGAGCTCTTGGAGCGAGCGAGCTGTTGTGA	1555
Qy	1582	GTGGGACCCGTTGGAGGAGAGTCTCCACAGGCTTCAGGGCTCTCTCTCCACGCGAAG	1641
Db	1556	GTGGGATCCCAAGGAGAGAGCTCCGCAAGCTTCAGGGCTGCTCTCTCATCGCAAG	1615
Qy	1642	GCATCTCAAACTCAATGCAAGTCTCCGCGACAGCTTTAGAGGCACTACCCCTAGCA	1701
Db	1616	GCATCTCAAACTCAATGGCAAGTCTTCCAGACAGCTTTGGAGCTCGCGGCCCAACA	1675
Qy	1702	CTTTGGCTCCCTGGACCAACTGGCTCTCTCCCATCTCTGAGCGCCGCGCCGCGCCCT	1761
Db	1676	CTTTGGCTCCCTGGATGAATCTGCCCACTCTGCCCTCTGGCCGCGCGAGCCCT	1735
Qy	1762	CAGGGCTGTGAGTGAAGCAGCATCTGTCTCTCGAGTCTTTGACCAATTTGACTTGC	1821
Db	1736	CAGGGCTGTGAGGAGGAGCAGCATCTGTCTCTCTGAGTCTTTGACCACTGACTTGC	1795
Qy	1822	CTGAAGCTCTTCCCGAAACCCCACTCAGGGGCTGTGTCTGTGGACAACTCAGGGGGC	1881
Db	1796	CTGAAGCTCTTCCAGAGCCCACTCGGGGCTGTGTGTGTGTGGACAACTCAGGGGGC	1855
Qy	1882	TTGAGCAGCCTCCTCTCAGAAG-----GTCTGAAGCGATGTGTGGAGAAATCCT	1929







AUTHORS Haferlach, T.; Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,  
Dugas, M., Eils, R., Brors, B. and Mergenthaler, S.  
TITLE Novel genetic markers for leukemias  
JOURNAL Patent: WO 03039443-A 2508 15-MAY-2003;  
Deutsches Krebsforschungszentrum (DE);  
Ludwig-Maximilian-Universitaet Muenchen (DE); Haferlach, Torsten,  
PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)  
FEATURES  
source 1.3443  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 47.4%; Score 1374.2; DB 6; Length 3443;  
Best Local Similarity 81.6%; Pred. No. 0;  
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;  
QY 82 CTACTGCTGCGGACCCACATCCACCTCGGGTCCCGGACCATGAGTTCGGTGCCTTAC 141  
DB 87 CTACTGATTCCTTCGCGGCCCTTGCTCACCTCTCTGCTCGCATGGAGTCGCTGTTTTG 146  
QY 142 TCAGAGCCCGGAGCCAGGCTCCCTCGGCTCCGCTCGGCTCGGAGAGCCCGCGCC 201  
DB 147 CGCGGCGCTCGGCGCCCACTCCCTCGGCGCGAGACTA-----GCCGCGCCG 194  
QY 202 TGGCGGACGGGCTCATCAAGTCGCTTAACTCTGTGTAAGAAGCAGCGCGTGAAGCGG 261  
DB 195 TGGCGGAAGGCTGATCAAGTCGCGCAAGCCCTTATGAAGAGCAGCGCGTGAAGCGG 254  
QY 262 ACCATCA CAACACAA CCTCGGCGACCGCTACGAGTTCTTGGAGACGCTGGGCAAGGGCA 321  
DB 255 ACCACCA CAAGCA CAACCTCGGCGACCGCTACGAGTTCTTGGAGACCTTGGGCAAGGGCA 314  
QY 322 CCTACGGGAAGGTGAAGAAGCAGAGAGCTCGGGGCTTGGTGGCCATCAAGTCCA 381  
DB 315 CCTACGGGAAGGTGAAGAAGCAGGAGCTCGGGGCGCTTGGTGGCCATCAAGTCAA 374  
QY 382 TCAGGAAGACAAATCAAGATGAGCAGGATCTGTGCAATACAGGAGGAGATTGAGA 441  
DB 375 TCGGAAGGACAAATCAAGATGAGCAGGATCTGTGCAATACAGGAGGAGATTGAGA 434  
QY 442 TCATGCTTCACTCAACACCCACATCATATGTCATCAAGTGTGTTGAGAATAGCA 501  
DB 435 TCATGTCATCACTCAACACCCCTCACATCATATGTCATCAAGTGTGTTGAGAATAGCA 494  
QY 502 GCAAGATTGTGATGTGATGGAGTATGCGAGCGGCGGATCTGATGATTACATCAGTG 561  
DB 495 GCAAGATCGTGATCGTATGGAGTATGCCAGCGCGGCGGACCTTTATGACTACATCAGCG 554  
QY 562 AGCGGCCACGCGTGAAGTGAAGGAGCGGAGGCAATTTCTCCGACAGATCGTGTGCGCC 621  
DB 555 AGCGGACGAGCTCAGTGAAGGCGAAGCTAGGCAATTTCTCCGCGAGATCGTCTGCGCG 614  
QY 622 TGCACTATGCGCACAGAGCGGATCGTTCAACGAGATCTCAAGCTGGAATAATCCTTC 681  
DB 615 TGCACTATGCGCATCAAGACAGAGTTGTCACACGAGATCTCAAGCTGGAATAATCCTTC 674  
QY 682 TAGATGCCAATGGAAACATCAAGATTGCTGACTTTGGCCCTTCCAACTGTACCAAAAG 741  
DB 675 TGGATGCCAATGGGAATATCAAGATTGCTGACTTTGGCCCTTCCAACTGTACCAATCAAG 734  
QY 742 GCAAGTTCTCCAGACGCTTGTGGAGCGCTCTCTACGCTCGCCCTGAGATAGTCAACG 801  
DB 735 GCAAGTTCTTGACAGCAATTTGTGGAGCGCCCTCTATGCTCGCCAGAGATTGTCAATG 794  
QY 802 GGAAGCCCTATGTGGGCCAGAGGTGAGCAGCTGGTCTCTGGGCGTTCTCTGTGTACATCC 861  
DB 795 GGAAGCCCTACACAGGCCACAGAGGTGAGCAGCTGGTCTCTGGGCGTTCTCTCTACATCC 854  
QY 862 TGGTGTGATGGACCATGCCCTTTGACCGGGAGGATCAATAAACA CTGGTGAAGCAATCA 921  
DB 855 TGGTGTGATGGACCATGCCCTTTGATGGGATGACCAATAAGATCTCTAGTGAAGACATCA 914

QY 922 GTAAACGGGGCTTACCGGTGAGCCGCCCAAGCGCTCCGATGCTGTGGCTGATCCGGTGGC 981  
DB 915 GCAACGGGGCTTACCGGGAGCCACCTAAACCTCTGATGCTGTGGCTGATCCGGTGGC 974  
QY 982 TGTAAATGTGTAACCCACCCGTCGGGCCACACATCGAGGATGTAGCCAGTCATTTGGTGG 1041  
DB 975 TGTGTGATGTGTAACCCACCCGTCGGGCCACCCCTGGAGGATGTGGCCAGTCACCTGTGG 1034  
QY 1042 TCAACTGGGGCTTACACACCGGAGTCGGGAAACAGGAAGCCCTGCGTGAAGGTGGGCACC 1101  
DB 1035 TCAACTGGGGCTTACGCCACCCGAGTGGAGAGCAGGAGGCTCCGCATGAGGGTGGGCACC 1094  
QY 1102 CTAGTGGTGAATTTGGCCGGGCTCCATGGCGGACTGGTTACGTCGCTCTCTCGCGCCCC 1161  
DB 1095 CTGGCAGTGAATCTGCTCGCGCGCTCCATGGCTGACTGGCTCGCGCGTCTCTCCCGCCCC 1154  
QY 1162 TCTGAGAAATGAGGCAAGGTGTGAGCTTCTTCAAGAGCAGTGTCCGGGAGGTGGAA 1221  
DB 1155 TCTTGAGAAATGGGGCCAAAGTGTGAGCTTCTTCAAGCAGCATGCACTTGGTGGGGAA 1214  
QY 1222 GCACTGTACTGGGCTGGAGCGGCAACATTTCTTAAAGAGTCCCGAAAGGAGAAATGACA 1281  
DB 1215 GCACCACTTGGGCTGGAGCGCCAGCATTCGCTCAAGAAAGTCCCGCAAGGAGAAATGACA 1274  
QY 1282 TGGCTCAAAATTCGAAGGTGACCCGGCTGAGGATACCTCTTCTCGCCCTGGCAAGAGCA 1341  
DB 1275 TGGCCAGTCTCTCCACAGTGACAACGGCTGATGACACTGCCCATCGCCCTGGCAGAGCA 1334  
QY 1342 GCCTTAAGCTTCCGAAAGGCAATTTCTCAAGAAAGTCTCTTA CTCTGTCAAGGGAGGTAC 1401  
DB 1335 ACCTCAAGCTGCCAAAGGGCAATTTCTCAAGAAAGGAGTGTGAGCTCTTGCAGAAAGGGGTAC 1394  
QY 1402 AGAGAGCCCTCAGGAAGTCAAGACCGGTGCTCATCTCCAGGGCAGCTGTCCCTGCTG 1461  
DB 1395 AGAGAGCCCTCAGGAGCTCAGCCCAATTCCTCGAGCCCGAGGGAGGCTGCCC----- 1448  
QY 1462 TATCCCTGCTCCCAAGGAAAGGCAATCTTAAAGAGTCTCGACAGCGTGAATCTGTTTACT 1521  
DB 1449 ---CGCTGCTCCCAAGAGGGCAATTTCTCAAGAAAGCCCGACAGCGAGTCTGGCTACT 1505  
QY 1522 ACTCTCTCCAGAGCCAGGAGTCTGGGGAACTTTTAGACGCGCAGTGATGTGTTGTGA 1581  
DB 1506 ACTCTCTCCAGAGCCAGTGAATCTGGGGAGCTCTTGGACGCGAGCGACGTGTTGTGA 1565  
QY 1582 GTGGGACCCCGTGGAGCAGAGTCTCCACAGGCTTCCAGGGCTCTCTCCACCGCAAG 1641  
DB 1566 GTGGGATCCCAAGGAGCAGAGCCCTCCGCAAGCTTTCAGGGCTGCTCTCCATCGCAAG 1625  
QY 1642 GCAATCTCAAACTCAATGGCAAGTTTCTCCGCAAGCTTTAGAGGCACTACCCCTTAGCA 1701  
DB 1626 GCATCTCAAACTCAATGGCAAGTTTCTCCAGACAGCTTGGAGCTCGCGGCCCCACCA 1685  
QY 1702 CTTTGGCTCTCGGACCAACTGGCTCTCTCCATCTCTGAGCCCGGCGCCAGCGCCCT 1761  
DB 1686 CTTTGGCTCTCGGATGAATCTCGCCCTCTCGCCCTTGGCGCCGCGCGAGCCCT 1745  
QY 1762 CAGGGCTGTGAGTGAAGCAGCATCTGTCTCCGAGTCTTTGACCAATTTGACTGCTG 1821  
DB 1746 CAGGGCTGTGAGGAGGACAGCATCTGTCTCTGAGTCTTTGAGTCTTTGAGTGTGCTG 1805  
QY 1822 CTGAAAGTCTTCCGGAACCCCACTGAGGGGCTGTGTGTCTGTGGACAACTGAGGGGG 1881  
DB 1806 CTGAAAGGCTCCAGAGCCCACTGCGGGCTGTGTGTCTGTGGACAACTGAGGGGG 1865  
QY 1882 TTGAGCAGCTCTCTCAGAAG-----GTCTGAAGCGATGTGTGACAGAAATCCT 1929  
DB 1866 TTGAGGAGCCCTCAGAGGGCCCTGGAAGCTGCTGAGGCGCTGCGCGCAGATCTCT 1925  
QY 1930 TGGGGATAGTCTCTTCTGACAGACTGCAAGAGGTGACTGCGAGCTTACAGACAG 1989  
DB 1926 TGGGGAGAGCTCTTTTCTGACAGACTGCGCAGGAGGTGACAGCGACCTACCGACAG 1985



```
Qy 1990 CCCTAGGAATCTGCTCAAAGCTCAGCTGAGGAGGAGATGGTCCCTAGTATGGGGTAG 2049
    |||
Db 1986 CACTGAGGCTGCTCAAAGCTCAGCTGAGTGGAGTAGGATGGCCAG-CCCGGTGAG 2044
    |||

Qy 2050 GCTCTGAGAGGTTTGCAGAGAAACCTTGGTGGATTCCTCC 2092
    |||
Db 2045 GCTCTCAGATCAGCTGTTGACCCCGGAGGAGATGCTTTC 2087
    |||

RESULT 15
HSM801859
LOCUS
DEFINITION Homo sapiens mRNA; cDNA DKFZp434J037 (from clone DKFZp434J037);
complete cds.
ACCESSION AL136891
VERSION AL136891.1 GI:12053280
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3443)
Koehler, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
Direct Submission
Submitted (12-MAR-2002) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by BMF (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
This clone (DKFZp434J037) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
Location/Qualifiers
1..3443
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="RZPD:DKFZp434J037"
    /db_xref="taxon:9606"
    /chromosome="1"
    /clone="DKFZp434J037"
    /tissue_type="testis"
    /clone_lib="434 (synonym: htes3). Vector pSport1; host
    DH10B; sites NotI + SalI"
    /dev_stage="adult"
1..3443
    /gene="DKFZp434J037"
128..2014
    /gene="DKFZp434J037"
    /note="strong similarity to serine/threonine-specific
    protein kinases"
    /codon_start=1
    /product="hypothetical protein"
    /protein_id="CAB66825.1"
    /db_xref="GI:12053281"
    /db_xref="GOA:Q9H093"
    /db_xref="TrEMBL:Q9H093"
    /translation="MESLVPARRSGTPPSAAELARPLAEGLIKSPKPLMKQAKRHH
    HKNLRRHYEFLTLGKTYGKVKARESSGRGLVYKISKDKIKDQDLMIHRRIE
    IMSSLNHPHTIAIHEVFENSSKIVIMEYASRGDLVDYISERQQLSEREARHFFROIV
    SAVHYCHQNRVVRDLKLENLLDANGNIKADFGLSNLYHQKFLQTCGSPLYASP
    EIVNGRPYGPVDSWLSGLVLLXILVHGTMPFDGDKILVQKISNGAREPPKPSDA
    CGILRWLLWMPFRTRLEVDVASHVWVNGYATRVGEQAPHEGHPGSDSARASWAD
    WLRRSRPLNENAKVCSFKKQVAGGSTTPGLERQSLKKSRKENDMAQSLHSDTA
    DDTAHPKSNLKLPGKILKGVASAGVQEDPELSPIPASPGQAAPLLKKGLIK
    KPRQESGYVSSPEPSGELLDPGVFVSGDPKEQPPQASGLLLHRKGLIKNKF
    SOTALALAAPTTGSLDEAPPLARAPSPGAVSDESLSSSEFDQLDLEPLPEP
    PLRGCVSVNLTSLBEPSPSGPSCSLRRWRQDPLGDSCLSDTCQEVATYRQALRVC
    SKLT"
3377..3382
    polyA_signal
```

```
polyA_site /gene="DKFZp434J037"
3399 /gene="DKFZp434J037"
ORIGIN
Query Match 47.4%; Score 1374.2; DB 9; Length 3443;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;
Qy 82 CTACTGCTGCCGACCCACCTCCACCTCGGGGTCCCGCCACCATGGAGTGGTGGCTTAC 141
    |||
Db 87 CTACTGATTCCTCGCGCCTTGCTCACTCTGCTGCGCATGGAGTGCCTGTTTCG 146
    |||
Qy 142 TCCAGCGCCGCGAGGCTCCCTCGCCCTCGCCCTCGGAGAGGCGCCGCGCCG 201
    |||
Db 147 CGCGGCGCTCCGGCCCACTCCCTCGCGCGCAGAGCTA-----GCCCGCGCGC 194
    |||
Qy 202 TGGCGGACGGGCTCATCAAGTCGCTTAAACCTCTGATGAAGACGAGGCGGTCAAGCGGC 261
    |||
Db 195 TGGCGGAGGGCTGATCAAGTCGCCCAAGCCCTAATGAAGAAGCAGGCGGTGAGCGGC 254
    |||
Qy 262 ACCATCAAAACAAACACCTCGCGGACCGCTACGAGTTCCTGGAGACGCTGGGCAAGGCA 321
    |||
Db 255 ACCACCACAAGCACAACTCGGCGCACCGCTACGAGTTCCTGGAGACCTTGGGCAAGGCA 314
    |||
Qy 322 CCTACGGGAGGTGAAGAAGGACGAGAGAGCTCGGGGGGTCTGGTGGCCATCAAGTCCA 381
    |||
Db 315 CCTACGGGAGGTGAAGAAGGCGCGGAGAGCTCGGGGGCGCTTGGTGGCCATCAAGTCAA 374
    |||
Qy 382 TCAGGAAAGACAAAATCAAGATGACGAGATCTGCTGCACATACGAGGAGGAGATTGAGA 441
    |||
Db 375 TCCGGAAGACAAAATCAAGATGACGAGATCTGATGCACATACGAGGAGGAGATTGAGA 434
    |||
Qy 442 TCATGCTTCACTCAACACCCCAACATCATTTGCCATTCATGAAGTGTGTTGAGAATAGCA 501
    |||
Db 435 TCATGTCATCACTCAACACCCCTCACATCATTTGCCATTCATGAAGTGTGTTGAGAACAGCA 494
    |||
Qy 502 GCAAGATTGTGATTGTCATGGAGTATGCCAGCGGAGCGGATCTGTATGATTACATCAGTG 561
    |||
Db 495 GCNAGATCGTGATCGTCATGGAGTATGCCAGCGGCGGAGACCTTTATGACTATCATCAGCG 554
    |||
Qy 562 AGCGGCCACGGCTGAGTGAAGCGGAGCGCGAGGATTTCTTCCGACAGATCGTGTCTGCCCC 621
    |||
Db 555 AGCGGCAGCAGCTCAGTGAGCGGAGCTAGGATTTCTTCCGCGAGATCGTCTCTGCGCG 614
    |||
Qy 622 TGCATCTATGCCACAGAAAGGAGTGTTCACGAGATCTCAAGCTGGAAGAAACATCCTTC 681
    |||
Db 615 TGCATCTATTGCCATCAGAAAGAGTTGTCCACCGGAGATCTCAAGCTGGAGAACATCCTCT 674
    |||
Qy 682 TAGATGCCAATGGAACATCAAGATTGTGACTTTTGGGCTCTCCACCTGTACCAAAAG 741
    |||
Db 675 TGGATGCCAATGGAAATATCAAGATTGTGACTTTGGGCTCTCCACCTCTACCATCAAG 734
    |||
Qy 742 GCAAGTTCCTCCAGAGCTTCTGTGGGAGCCCTCTCTACGCGCTCGCTGAGATAGTCAACG 801
    |||
Db 735 GCAAGTTCCTGCAGACATTTCTGTGGGAGCCCTCTCTATGCTCGCCGAGAGATTGTCAATG 794
    |||
Qy 802 GGAAGCCCTATGTGGGCCCGAGAGGTGGACAGCTGCTCTCTGGGCGTTCTCTGTATCATCC 861
    |||
Db 795 GGAAGCCCTATACAGAGCCCGAGAGGTGGACAGCTGGTCCCTCGGCTCTCTCTCTATCATCC 854
    |||
Qy 862 TGGTGCATCGCACCATGCCCTTTTGAAGCGGACAGATCATAAAAACACTGGTGAAGCAATCA 921
    |||
Db 855 TGGTGCATCGCACCATGCCCTTTTGAAGCGGACAGATCATAAAAACACTGGTGAAGCAATCA 914
    |||
Qy 922 GTAAAGCGGCTTACCGTGAAGCGCCCGCAAGCCGTCGAGTCCGATCGCTGTGAGTCCGCTGCG 981
    |||
Db 915 GCAAGCGGCGCTTACCGGAGGACCACTTAAACCTCTGTATGATGCTGTGCGCTGTATCCGCTGCG 974
    |||
Qy 982 TGTTAATGTTGAACCCCGCTCGGGGCCACACTCGAGGATGTAGCCAGTCACTTGTGTGG 1041
    |||
Db 975 TGTGATGTTGAACCCCGCTCGGGGCCACCTCGAGGATGTGGCCAGTCACTGTGTGG 1034
    |||
```



Search completed: January 25, 2005, 01:53:15  
Job time : 12193.3 secs

```
Qy 1042 TCAACTGGGTTACACACCGAGTCGGGGAACAGGAAGCCCTGCTGAGGGTGGGCACC 1101
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1035 TCAACTGGGCTACGCCACCGAGTCGGGAGAGCAGGAGGCTCCGCATGAGGGTGGGCACC 1094
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1102 CTAGTGGTGAATTTGGCGGGGCTCCATGCGGAGCTGGTTACGTGCTCTCTCGGCCCCC 1161
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1095 CTGCACTGACTCTGCGCGGCTCCATGCGTGAATGCTGCTCGGCGTTCTCCGCCCCC 1154
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1162 TCCTGGAGAATGGAGCCAAAGTGTGCAGCTTCTTCAAGCAGCAGCTGCGCGGAGGTGAA 1221
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1155 TCCTGGAGAATGGGCGCAAGTGTGCAGCTTCTTCAAGCAGCAGCTGCTGGTGGGGAA 1214
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1222 GCATCTGACTCTGGGCTGGAGCGGCAACATCTCTTAAAGAGTCCGAAAGGAGATGACA 1281
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1215 GCACCACTCTGGCTGGAGCGCCAGCATTCGCTCAAGAGTCCCGCAAGAGATGACA 1274
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1282 TGGCTCAAAATCTGCAAGGTGACCGGCTGAGGATACCTCTTCTCGGCCCTGGCAAGACA 1341
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1275 TGGCCCACTCTCTCCACAGTGACACGGCTGATGACACTGCCCCATCGCCCTGGCAAGACA 1334
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1342 GCCTTAAAGCTTCCGAAAGGCAATCTCAAGAAAGTCTCTTACCTCTCAGGGAGGTAC 1401
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1335 ACCTCAAGCTGCCAAGGGCAATCTCAAGAAAGGTGTGAGCTCTGCAAGGGGTAC 1394
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1402 AGGAGGACCTCAAGAACTCAGACCGGTGCTGATACCTCCAGGGCAGCCTGTCCCTGCTG 1461
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1395 AGGAGGACCTCCGAGCTCAGCCCAATCCCTCGAGCCCGAGGGCAGGCTGCC----- 1448
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1462 TATCCCTGCTCCCAAGAAAGGCAATCTTAAAGTCTCGACAGCGTGAATCTGGTTACT 1521
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1449 ---CGTGTCTCCCAAGAAAGGCAATCTCAAGAAAGCCCGACAGCGGAGTCTGGCTACT 1505
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1522 ACTCCTCTCCAGAGCCAGCGAGTCTGGGAACTCTTAGAGCCGCGAGTGATGTGTTGTGA 1581
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1506 ACTCCTCTCCGAGCCAGCGAGTGAATCTGGGAGCTCTTGGAGCGAGGCGACGTGTTGTGA 1565
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1582 GTGGGAGCCCGTGGAGCAGAACTCTCCAGGCTTCAGGGCTCTCTCTCCACCGCAAGG 1641
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1566 GTGGGATCCCAAGGAGCAGAGCTCCGCAAGCTTCAGGGCTGCTCTCCATCGCAAG 1625
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1642 GCATTTCAAACTCAATGGCAAGTCTCCGCAAGCTTTAGAGGCACTACCCCTAGCA 1701
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1626 GCATCTCAAACTCAATGGCAAGTCTCCAGAGCTTTGGAGCTCGCGGCCCCCA 1685
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1702 CTTTGGCTCCCTGGACCACTGGCTCTCTCCATCTCGAGCCCGCCGCGCCCT 1761
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1686 CTTTGGCTCCCTGGATGAATCGCCCACTCGCCCTCGGCCCGGCGAGCCCT 1745
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1762 CAGGGCTGTGAGTGAGGACAGCATCTGTCTCCGAGTCTTTGACCAATTGACTTGC 1821
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1746 CAGGGCTGTGAGCGAGGACAGCATCTGTCTCTGAGTCTTTGACCACTGACTTGC 1805
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1822 CTGAACGTCTTCCGAAACCCCACTGAGGGGCTGTGTCTGTGGACAACCTGAGGGGC 1881
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1806 CTGAACGGCTCCAGAGCCCCCACTGCGGGGCTGTGTGTGTGGACAACCTCACGGGC 1865
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1882 TTGAGCAGCTCTCCCTCAGAAG-----GTCTGAAGCGATGTGTGCGAGATCTCT 1929
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1866 TTGAGGAGCCCCCTCAGAGGGCCCTGGAAGCTGCTGAGGCGCTGGCGGAGATCTCT 1925
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1930 TGGGGGATAGTGTCTTTCTGTGACAGACTGCCAAGAGGTGACTGCAAGCCTACAGACAAG 1989
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1926 TGGGGGACAGTGTCTTTCTGTACAGACTGCCAGGAGGTGACAGCGACCTACCGACAGG 1985
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1990 CCCTAGGAATCTGCTCAAGCTAGCTGAGGAGGGAGATGGTCCCTAGTATGGGGTAG 2049
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1986 CACTGAGGGTCTGCTCAAGCTCACCCTGAGTGGAGTAGGCAATTCGCCAG-CCCGGTGAG 2044
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2050 GCTCTGAGAGGGTTTGAGAGGAACCTGTGCTGCTGCTCC 2092
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2045 GCTCTGAGATGAGCTGCTGTTGACCCCGAGGGAGATGCTCTTC 2087
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 05:56:58 ; Search time 1342.36 Seconds  
(without alignments)  
11348.500 Million cell updates/sec

Title: US-09-980-464-4

Perfect score: 2902

Sequence: 1 cactagtgcattcaagaat.....aaaaaaaaaaaaaaaaaaaaa 2302

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_23Sep04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2902	100.0	2902	4 AAC90433	Aac90433 Murine Ly
2	2066.8	71.2	2929	6 AAD31710	Aad31710 Rat SNF1/
3	1379	47.5	3353	6 ABK14000	Abk14000 cDNA enco
4	1379	47.5	3353	12 ADL14160	Adl14160 Novel hum
5	1375.8	47.4	3395	4 AAK94280	Aak94280 Human ful
6	1375.8	47.4	3395	12 ADL30885	Adl30885 Full leng
7	1374.2	47.4	3443	5 ABX71420	Abx71420 Human cel
8	1374.2	47.4	3443	10 ADF76964	Adf76964 Novel hum
9	1374.2	47.4	3443	10 ADF81952	Adf81952 Leukaemia
10	1374.2	47.4	3443	12 ADO20171	Ado20171 Human PRO
11	1374.2	47.4	3463	12 ADJ96554	Adj96554 Human cal
12	1372.2	47.3	2501	12 ADM43851	Adm43851 Novel hum
13	1371.8	47.3	2043	6 ABZ11333	Abz11333 Human pol
14	1370.6	47.2	3360	6 AAD26459	Aad26459 Human kin
15	1360.2	46.9	1884	12 ADL14162	Adl14162 Novel hum
16	1349.2	46.5	2291	4 AAF44659	Aaf44659 Novel pro
17	1349.2	46.5	2291	12 ADI29357	Adi29357 Human-MAR
18	1323.2	45.6	3200	4 AAF75338	Aaf75338 Human TGF
19	1230	42.4	1833	10 ABZ77163	Abz77163 Human pro
20	849.2	29.3	2616	6 ABQ72599	Abq72599 Human MDD
21	849.2	29.3	2619	6 ABQ72698	Abq72698 Human MDD

22	507.6	17.5	6828	4 AAD14327	Aad14327 Human pro
23	507.6	17.5	6828	10 ADE38420	Ade38420 Human pro
24	507.6	17.5	6828	11 ADN95767	Adn95767 Human BEC
25	507.6	17.5	6828	12 ADJ74808	Adj74808 Marker ge
26	507.6	17.5	6828	12 ADL25352	Adl25352 Human ARK
27	507.6	17.5	6828	12 ADQ19733	Adq19733 Human sof
28	507.6	17.5	6828	12 ADP43253	Adp43253 Human sof
29	507.6	17.5	6854	12 ADQ23883	Adq23883 Human sof
30	506.4	17.5	2884	4 AAD14328	Aad14328 Human pro
31	360.2	12.4	587	4 AAK93296	Aak93296 Human CDN
32	360.2	12.4	587	4 AAK91887	Aak91887 Human CDN
33	360.2	12.4	587	12 ADL29723	Adl29723 5' end of
34	360.2	12.4	587	12 ADL28314	Adl28314 5' end of
35	356.8	12.3	3594	5 AAD03994	Aad03994 Human pro
36	296.8	10.2	1454	4 ABL21269	AbL21269 Drosophil
37	268.2	9.2	1723	4 AAH13802	Aah13802 Human CDN
38	236.6	8.2	1594	4 AAF44655	Aaf44655 Novel pro
39	236.6	8.2	1594	8 AAL60326	Aal60326 Human 207
40	236.6	8.2	1594	12 ADI29353	Adi29353 Mouse MAR
41	236.6	8.2	3170	6 ABA05737	Abao5737 Murine ne
42	236.6	8.2	3250	6 ABA05738	Abao5738 Murine ne
43	233.6	8.0	1549	6 ABS64386	Abbs64386 Human ser
44	233.6	8.0	2085	8 AAD51409	Aad51409 Human mic
45	233.6	8.0	2085	10 ADG91747	Adg91747 Human mic

#### ALIGNMENTS

RESULT 1  
AAC90433  
ID AAC90433 standard; cDNA; 2902 BP.

AC AAC90433;

DT 19-MAR-2001 (first entry)

DE Murine Lymph node Stromal cell kinase 1 coding sequence.

XX Murine; Lymph node Stromal cell kinase; MLSK-1; autoimmune disorder;  
XX wound healing; periodontal disease; inflammatory disease; tumour;  
XX infection; allergy; ss.

OS Mus musculus.

PN WO2000073468-A1.

PD 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US014696.

PF 28-MAY-1999; 99US-0136781P.

PR (IMMV ) IMMUNEX CORP.

XX Bird TA, Virca GD, Martin U, Anderson DM;

DR WPI; 2001-061546/07.

DR P-PSDB; AAB50056.

XX Novel murine and human kinase nucleic acids useful for treating  
XX inflammations, infections, tumors, allergies, autoimmune diseases, and  
XX for stimulating or suppressing immune responses.

PS Claim 1; Page 86-87; 106pp; English.

XX The present sequence is the coding sequence for Murine Lymph node Stromal  
XX cell kinase 1 (MLSK-1). The protein encoded by the present sequence is  
XX useful for treating a variety of disorders listed in the disclosure of  
XX the specification, including autoimmune disorders, allergic reactions,  
XX myeloid or lymphoid cell deficiencies, wound healing and tissue repair  
XX and replacement, burns, incisions and ulcers, periodontal disease,  
XX inflammatory diseases, tumours and bacterial, viral or fungal infection

XX	Sequence	2902 BP; 654 A; 794 C; 796 G; 658 T; 0 U; 0 Other;
SQ	Query Match	100.0%; Score 2902; DB 4; Length 2902;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 2902; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	CAC TAGTGGATCCAAAGAAATTCGACGAGCGCTGCTCGGTGCGGTGTGACCTCTGAG 60
Db	1	CAC TAGTGGATCCAAAGAAATTCGACGAGCGGTGCTCGGTGCGGTGTGACCTCTGAG 60
Qy	61	CCC GCGGCTCAGCGCGCTGCTACTGCTGCCCGACCCACCTCCACCTCGCGGTCCCGCA 120
Db	61	CCC GCGGCTCAGCGCGCTGCTACTGCTGCCCGACCCACCTCCACCTCGCGGTCCCGCA 120
Qy	121	CCATGGAGTGGTGGCTTACTCAGCGCCCGAGCGAGCTCCCTCGGCTCCGCGCTGG 180
Db	121	CCATGGAGTGGTGGCTTACTCAGCGCCCGAGCGAGCTCCCTCGGCTCCGCGCTGG 180
Qy	181	CCTCGGAGAGCGCCGCGCTGCGGACGGCTCATCAAGTCCCTAAACCTCTGATGA 240
Db	181	CCTCGGAGAGCGCCGCGCTGCGGACGGCTCATCAAGTCCCTAAACCTCTGATGA 240
Qy	241	AGAAAGCGGGTGAAGCGGCACCATCAAAACAACCTCGCGCACCGCTACGAGTTC 300
Db	241	AGAAAGCGGGTGAAGCGGCACCATCAAAACAACCTCGCGCACCGCTACGAGTTC 300
Qy	301	TGGAGACGCTGGCAAGGCGACCTACCGGAGGTGAAGAGGACGAGAGCTCGGGG 360
Db	301	TGGAGACGCTGGCAAGGCGACCTACCGGAGGTGAAGAGGACGAGAGCTCGGGG 360
Qy	361	GTCTGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGCAGGATCTGTGC 420
Db	361	GTCTGTGGCCATCAAGTCCATCAGGAAGACAAATCAAGATGAGCAGGATCTGTGC 420
Qy	421	ACATACGAGGAGATGAGATCATGCTTCACTCAACCCACCATCATGTCATCC 480
Db	421	ACATACGAGGAGATGAGATCATGCTTCACTCAACCCACCATCATGTCATCC 480
Qy	481	ATGAAGTGTGAGATGAGCAGGAGATGATGATGTCGAGTATGCGAGGAGCG 540
Db	481	ATGAAGTGTGAGATGAGCAGGAGATGATGATGTCGAGTATGCGAGGAGCG 540
Qy	541	ATCTGTATGATTAATCAGTGGCGGCGGCTGAGTGGAGCGGAGCCAGGCAATTC 600
Db	541	ATCTGTATGATTAATCAGTGGCGGCGGCTGAGTGGAGCGGAGCCAGGCAATTC 600
Qy	601	TCCGACAGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db	601	TCCGACAGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy	661	TCAAGCTGGAACATCCTTCTAGATGCGCAATGGAACATCAAGATGCTGCTTGGCC 720
Db	661	TCAAGCTGGAACATCCTTCTAGATGCGCAATGGAACATCAAGATGCTGCTTGGCC 720
Qy	721	TCTCCACCTGTATACCAAGGCAAGTTCCTCCAGAGCTTCTGAGGAGCCCTCTACG 780
Db	721	TCTCCACCTGTATACCAAGGCAAGTTCCTCCAGAGCTTCTGAGGAGCCCTCTACG 780
Qy	781	CCTGCGCTGAGATGATCAAGGAGAGCCCTATGTTGGGCGCAGAGGTGGAAGTGTTC 840
Db	781	CCTGCGCTGAGATGATCAAGGAGAGCCCTATGTTGGGCGCAGAGGTGGAAGTGTTC 840
Qy	841	TGGGCGTTCCTGTATACCTCTGTGATGCGACCATGCGCTTGGGCGCAGGATCAT 900
Db	841	TGGGCGTTCCTGTATACCTCTGTGATGCGACCATGCGCTTGGGCGCAGGATCAT 900
Qy	901	AAACACTGTGTAAGCAAAATCAGTAACGGGCTTACCGTGGAGCGCCCAAGCGCTG 960
Db	901	AAACACTGTGTAAGCAAAATCAGTAACGGGCTTACCGTGGAGCGCCCAAGCGCTG 960
Qy	961	CCTGTGGCTGATCCGGTGGCTGTAAATGTGTGAACCCACCGCTCGGCGCACATGAGG 1020
Db	961	CCTGTGGCTGATCCGGTGGCTGTAAATGTGTGAACCCACCGCTCGGCGCACATGAGG 1020

Db	961	CCTGTGGCTGATCCGGTGGCTGTAAATGTGTGAACCCACCGCTCGGCGCACATGAGG 1020
Qy	1021	ATGTAGCCAGTCAATGTTGGTGGTCAACTGGGTTTACACCCGAGTGGGGAACAGGAAG 1080
Db	1021	ATGTAGCCAGTCAATGTTGGTGGTCAACTGGGTTTACACCCGAGTGGGGAACAGGAAG 1080
Qy	1081	CCCTGGTGGAGTGGGCAACCTAGTGGTGGTGGTGGGCGGCTCCATGGCGGACTGGT 1140
Db	1081	CCCTGGTGGAGTGGGCAACCTAGTGGTGGTGGTGGGCGGCTCCATGGCGGACTGGT 1140
Qy	1141	TACGTGCTCCTCGCGCCCTCCTCGAGAAATGAGCAAGGTGTGAGCTTCTTCAAGC 1200
Db	1141	TACGTGCTCCTCGCGCCCTCCTCGAGAAATGAGCAAGGTGTGAGCTTCTTCAAGC 1200
Qy	1201	AGACAGTGGCGGAGTGGAGCACTGTACTCTGGCTGGAGCGGCAACATTTCTTAAAG 1260
Db	1201	AGACAGTGGCGGAGTGGAGCACTGTACTCTGGCTGGAGCGGCAACATTTCTTAAAG 1260
Qy	1261	AGTCCCGAAGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCT 1320
Db	1261	AGTCCCGAAGAGAAATGACATGGCTCAAAATCTGCAAGGTGACCCGGCTGAGGATACCT 1320
Qy	1321	CTTCTCGCTCGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATTTCAAGAAAAGTCT 1380
Db	1321	CTTCTCGCTCGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATTTCAAGAAAAGTCT 1380
Qy	1381	CTACCTCGTCAGGGAGGTACAGGAGACCTCAGAACTCAGACCGGTGCTGATCTC 1440
Db	1381	CTACCTCGTCAGGGAGGTACAGGAGACCTCAGAACTCAGACCGGTGCTGATCTC 1440
Qy	1441	CAGGCGAGCTGTCTCTGTCTGATCTCTCTCCCAAGGAAGCATCTTAAAGAGTCT 1500
Db	1441	CAGGCGAGCTGTCTCTGTCTGATCTCTCTCCCAAGGAAGCATCTTAAAGAGTCT 1500
Qy	1501	GACAGGTGAATCTGTTTACTCTCTCTCAGAGCCAGCGAGTCTGGGGAATCTCTAG 1560
Db	1501	GACAGGTGAATCTGTTTACTCTCTCTCAGAGCCAGCGAGTCTGGGGAATCTCTAG 1560
Qy	1561	ACGCCAGTGTGTTTGTGAGTGGGAGCCCGTGGAGCAGAGTCTCCAGGCTTCAG 1620
Db	1561	ACGCCAGTGTGTTTGTGAGTGGGAGCCCGTGGAGCAGAGTCTCCAGGCTTCAG 1620
Qy	1621	GGCTCTCTCTCCACCCCAAGGGCAATTTCAAACTCAATGGCAAGTTCCTCCGACAGCCT 1680
Db	1621	GGCTCTCTCTCCACCCCAAGGGCAATTTCAAACTCAATGGCAAGTTCCTCCGACAGCCT 1680
Qy	1681	TAGAAGGCACTACCCCTAGACACTTTTGGCTCTCTGAGCAAACTGGCTCTCTCCATCTG 1740
Db	1681	TAGAAGGCACTACCCCTAGACACTTTTGGCTCTCTGAGCAAACTGGCTCTCTCCATCTG 1740
Qy	1741	CAGCCCGGCGGCGGCTCAGGGCTGTGAGTGGAGCAGCATCTCTGCTCTCGAGT 1800
Db	1741	CAGCCCGGCGGCGGCTCAGGGCTGTGAGTGGAGCAGCATCTCTGCTCTCGAGT 1800
Qy	1801	CTTTTGACCAATGGCAATTTGCTGAACTTCTTCCGAAACCCCACTGAGGGGTGTGTGT 1860
Db	1801	CTTTTGACCAATGGCAATTTGCTGAACTTCTTCCGAAACCCCACTGAGGGGTGTGTGT 1860
Qy	1861	CTGTGCAAACTGAGGGGCTTGGAGCGCTCTCTCAGAAAGTCTGAAAGCGATGTGTGC 1920
Db	1861	CTGTGCAAACTGAGGGGCTTGGAGCGCTCTCTCAGAAAGTCTGAAAGCGATGTGTGC 1920
Qy	1921	AGGAATCTTTGGGAGTGTCTGCTGAGCTGAGGAGGAGTGTGTGTGTGTGTGTGT 1980
Db	1921	AGGAATCTTTGGGAGTGTCTGCTGAGCTGAGGAGGAGTGTGTGTGTGTGTGTGT 1980
Qy	1981	ACAGCAAGCCCTAGGAATCTGCTCAAGCTCAGCTGAGGAGGAGAGTGTGTGTGTGT 2040
Db	1981	ACAGCAAGCCCTAGGAATCTGCTCAAGCTCAGCTGAGGAGGAGAGTGTGTGTGTGT 2040
Qy	2041	ATGGGGTAGGCTCTGAGAGGTTTGGAGGAAACCTTGGGTGGATTTCTCCAGTGAATA 2100
Db	2041	ATGGGGTAGGCTCTGAGAGGTTTGGAGGAAACCTTGGGTGGATTTCTCCAGTGAATA 2100

```
Qy 2101 GAGTACATCAAGGGCTCTAGCTCTGACGCTGACTGAACCTGAAAGATGAGAAATCGC 2160
Db 2101 GAGTACATCAAGGGCTCTAGCTCTGACGCTGACTGAACCTGAAAGATGAGAAATCGC 2160
Qy 2161 ATTGATGTGAAAGGAATGGAAACCCCTGTGCTGCCGAGTGTATAGTGGGGTGGCCTGAA 2220
Db 2161 ATTGATGTGAAAGGAATGGAAACCCCTGTGCTGCCGAGTGTATAGTGGGGTGGCCTGAA 2220
Qy 2221 GGTGCTTACCTCTCTTGTGCGCATGAGTGTACCCATGACATTTCCACCCCTGTCTCTGG 2280
Db 2221 GGTGCTTACCTCTCTTGTGCGCATGAGTGTACCCATGACATTTCCACCCCTGTCTCTGG 2280
Qy 2281 CTGCACCTTCACATAAGTTCTTGTGTTCCATCAACACCCAGGGTTAGAACCTCGACTTCCT 2340
Db 2281 CTGCACCTTCACATAAGTTCTTGTGTTCCATCAACACCCAGGGTTAGAACCTCGACTTCCT 2340
Qy 2341 GGGAGGTAAATGTAGTACTGCCATTAATTTAGAGAGAAACAGCCTCTGGTTTCCATCT 2400
Db 2341 GGGAGGTAAATGTAGTACTGCCATTAATTTAGAGAGAAACAGCCTCTGGTTTCCATCT 2400
Qy 2401 CTGCTGCTGTGCATCTCAAGAGACTGGGAAGACTCGGACCGCTGTTTGACTTCTATCTCA 2460
Db 2401 CTGCTGCTGTGCATCTCAAGAGACTGGGAAGACTCGGACCGCTGTTTGACTTCTATCTCA 2460
Qy 2461 GGGGACAGATGCCCTCGGACCCCATCTTAGATCTCAGAGACTTGAACCTTGAAGCTGTT 2520
Db 2461 GGGGACAGATGCCCTCGGACCCCATCTTAGATCTCAGAGACTTGAACCTTGAAGCTGTT 2520
Qy 2521 CCTAGTACCAGATGTGGATGGATGCTCTGTTTCTCAGGCCAACGGGACCTAGAAATGTC 2580
Db 2521 CCTAGTACCAGATGTGGATGGATGCTCTGTTTCTCAGGCCAACGGGACCTAGAAATGTC 2580
Qy 2581 TGACTTATTTATTTTGTGATCTCAGCTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 2640
Db 2581 TGACTTATTTATTTTGTGATCTCAGCTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 2640
Qy 2641 TTTTGTGTTTAAAGTGAATTTTGTGCTTTCAATGAATGTGAATGCTGTGTTCTGGGGAAC 2700
Db 2641 TTTTGTGTTTAAAGTGAATTTTGTGCTTTCAATGAATGTGAATGCTGTGTTCTGGGGAAC 2700
Qy 2701 TCACCTGTGCTGCTGAAAGTTATGTACAGAGATATTTGGCAATGATGTCCTCTATTC 2760
Db 2701 TCACCTGTGCTGCTGAAAGTTATGTACAGAGATATTTGGCAATGATGTCCTCTATTC 2760
Qy 2761 AAGGGGGTGGGGCGTTTTCAAATGATGTCTTGAGCACTGCTCGATTGATCTCCA 2820
Db 2761 AAGGGGGTGGGGCGTTTTCAAATGATGTCTTGAGCACTGCTCGATTGATCTCCA 2820
Qy 2821 GTCCCTTCACACCCCAAGGCTGGCCACCCTCCCTCATCTTCTGTCGCAAAAAA 2880
Db 2821 GTCCCTTCACACCCCAAGGCTGGCCACCCTCCCTCATCTTCTGTCGCAAAAAA 2880
Qy 2881 AAAAAAAAAAAAAAAAAAAAAA 2902
Db 2881 AAAAAAAAAAAAAAAAAAAAAA 2902

RESULT 2
AAD31710
ID AAD31710 standard; cdna; 2929 BP.
XX
AC AAD31710;
XX
XX
DT 18-JUN-2002 (first entry)
DE
DE DE
XX
XX Rat SNF1/AMPK-Related Kinase (SNARK) cDNA.
KW Rat; SNF1/AMPK-Related Kinase; SNARK; enzyme; stress response; diabetes;
KW glucose deprivation; lipid metabolism; therapy; lipoprotein disorder;
KW hyperglycaemic; drug screening; hypoglycaemia; ss.
XX
OS Rattus sp.
```

```
XX Key Location/Qualifiers
PH 83..1975
FT /tag= a
FT /product= "Rat SNARK protein"
XX
PN WO200212456-A2.
XX 14-FEB-2002.
XX
XX 02-AUG-2001; 2001WO-CA001109.
XX
XX 03-AUG-2000; 2000US-0222650P.
XX 12-MAR-2001; 2001US-0274613P.
XX 28-MAR-2001; 2001CA-02340783.
XX
XX (ONEO-) 1149336 ONTARIO INC.
XX
XX Drucker DJ, Rosen CF, Lefebvre DL;
XX WPI; 2002-241747/29.
XX P-PSDB; AAE19885.
XX
XX AMPK (AMP-activated protein kinase)-related kinase, designated SNARK
polypeptides and polynucleotides, useful for treating or preventing
diabetes, or other disorders of lipoprotein production leading to
increased levels of cholesterol.
XX
XX Example; Fig 2; 94pp; English.
XX
XX The invention relates to an AMPK (AMP-activated protein kinase)-related
kinase, designated SNARK polypeptides and polynucleotides. SNARK
(SNF1/AMP-activated protein kinase) is involved in stress response to
glucose deprivation. The polynucleotides are useful for expressing SNARK
protein in isolated form or as a protein conjugate. Activation of SNARK
stimulates liver CPT-1 thus enhances lipid metabolism in liver cells and
in other cell types such as heart and skeletal muscles, as well as
increases GLUT-4 and glycogen in muscle. Activation of SNARK is predicted
to have insulin-like effects that would enhance the disposal of glucose
into muscle and reduce plasma glucose for the treatment of diabetes and
some type of disorders of lipoprotein production leading to increased
levels of cholesterol or triglycerides. SNARK or its variants may be
administered to a subject to treat or prevent a disease associated with
decreased expression of SNARK, such as diabetes. SNARK antibodies are
used to modulate SNARK activity either in vivo for therapeutic purposes,
or in vitro, for drug screening and related investigational purposes.
XX SNARK antagonists may be administered to increase fuel production,
decrease glucose uptake and increase levels of blood glucose in a patient
suffering from hypoglycaemia. The present sequence is rat SNARK cDNA
XX
SQ Sequence 2929 BP; 650 A; 788 C; 791 G; 700 T; 0 U; 0 Other;
Query Match 71.2%; Score 2066.8; DB 6; Length 2929;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 2488; Conservative 0; Mismatches 277; Indels 115; Gaps 14;
Qy 49 GTGACCTCTGAGCCCGCGCTCAGCGCTGCTACTGCTGCCGCCACCTCCACCTC 108
Db 9 GTGACCTCTGAGCTCGGCTCTCCGCGCTGCTGCTGCCGCCACCTCCCGCTC 68
Qy 109 GCGGTCCCGCACCATGGAGTGGTGGCTTACTCCAGCGCCCGAGCCAGGCTCCCTCGG 168
Db 69 GCGGTCCCGCACCATGGAGTGGTGGCTTACTACCCGCGCGGAACTGGCTCCCTCGG 128
Qy 169 CTTCCGCTTGGCTCGGAGAGCGCCGCGCTGGCGAGCGGCTCATCAAGTCGCTA 228
Db 129 CTTCCGCTTGGCTCGGAGAGCGCCGCGCTGGCGAGCGGCTCATCAAGTCGCTA 188
Qy 229 AACCTCTGATGAAGAAGAGCGGCTGAAGCGGACCATCAACAACCACTCGCGCAC 288
Db 189 AACCTCTGATGAAGAAGAGCGGCTGAAGCGGACCATCAACAACCACTGAGGAC 248
Qy 289 GCTACGAGTTCTTGGAGACGCTGGGCAAGGSCACCTACGGGAAGGTGAAGAAGCAG 348
```

Db 249 GCTACGAGTCTCTGGAGACCTTGGCAGAGGCACTACGGGAGGTTGAGAAAGCAGCAG 308  
Qy 349 AGAGCTCGGGCGGTCTGTGGCCATCAAGTCCATCAGGAAGAAGCAAAATCAAGATGAGC 408  
Db 309 AGAGCTCGGGAGCGCTGTGGCCATCAAGTCTATCAGGAAGGAAGAAATCAAGATGAGC 368  
Qy 409 AGGATCTGCTCACATACGAGAGGAGATTGAGATCATGTCTTCACTCAACACACCCCA 468  
Db 369 AGGATCTGTTGCACATAAGAGAGGAGATCGAGATCATGTCTTCACTCAACACACCCCA 428  
Qy 469 TCATTGCCATCCATGAAGTGTGAGATAGCAGCAAGATTGTGATTGTCTATGAGTATG 528  
Db 429 TCATTGCCATCCATGAAGTGTGAGAACAGCAGCAAGATTGTGATTGTCTATGAGTATG 488  
Qy 529 CCAGCCGAGGCGATCTGTATGATTACATCATGAGCGGCCACCGCTGAGTGAAGCGGAGC 588  
Db 489 CCAGCCGAGGCGATCTGTACGATTACATCATGAGCGGCCACCGCTGAATGAGCGGAGC 548  
Qy 589 CCAGGCAATTTCTCCGACAGATCGTGTCTGCCCTGCACTACTGCGCACCAAGACCGGATCG 648  
Db 549 CCAGGCAATTTCTCCGACAGATCGTGTCCGCCCTGCACTACTGCGCACCAAGACCGGATG 608  
Qy 649 TTACCCGAGATCTCAAGCTGGAACAATCTCTTAGATGCCAATGGAACAATCAAGATTG 708  
Db 609 TTACCCGAGATCTCAAGCTGGAACAATCTCTTAGATGCCAATGGAACAATCAAGATTG 668  
Qy 709 CTGACTTTGGCTCTCCAACTGTACCAAAAGGAAAGTTCTCTCAGACGTTCTGTGGGA 768  
Db 669 CTGATTTTGGCTCTCCAACTGTATCAAAAGGCAAGTTCTCTCAGACGTTCTGTGGGA 728  
Qy 769 GCCCTCTACGCTCGCTGAGATAGTCAACGGGAGCCCTATGTGGGCCAGAGGTGG 828  
Db 729 GCCCTCTATGCTCTACCTGAGATCGTCAACGGGAGCCCTATGTGGGCCAGAGGTGG 788  
Qy 829 ACAGCTGGTCTCTGGCGGTCTCTCTGATACATCTCTGTGATGCAATGCCATGCCCTTTGACG 888  
Db 789 ACAGCTGGTCTCTGGCGGTCTCTCTGATACATCTCTGTGATGCAATGCCATGCCCTTTGACG 848  
Qy 889 GGCAGGATCATAAACACTGTGTGAAGCAAAATCAGTAAAGCGGCTTACGTTAGCGGCCCA 948  
Db 849 GGCAGGATCATAAACACTGTGTGAAGCAAAATCAGTAAAGCGGCTTACCGAGAGCCGTGCA 908  
Qy 949 AGCCGTCCGATGCTGTGGCTGATCGGTGGCTGTTAATGTGAACCCCAACCGTCGGG 1008  
Db 909 AACGCTCTGATGCTGTGGCTGATCGGTGGCTGTTAATGTGAACCCCAACCGTCGGG 968  
Qy 1009 CCACACTGGAGGATGTAGCCAGTCAATTTGGTGGTCAACTGGGTTACAGCACCCGATTCG 1068  
Db 969 CCACACTGGAGGATGTAGCCAGTCAATTTGGTGGTCAACTGGGTTACAGCACCCGATTCG 1028  
Qy 1069 GGGNACAGGAAGCCCTGGTGGAGGTTGGGACCCCTAGTGTGACTTTGGCCGGGCTTCCA 1128  
Db 1029 GGGNACAGGAAGCTCTCGAGAGGTTGGGACCCCTAGCGTGAATCTGGCCGGGCTCTTA 1088  
Qy 1129 TGGCGGACTGTTACGTGCTCTCTCGCGCCCTCTCTGGAGAAATGGAGCCAAAGTGTGCA 1188  
Db 1089 TGGCGGACTGTTACGTGCTCTCTCGCGCCCTCTCTGGAGAAATGGAGCCAAAGTGTGTA 1148  
Qy 1189 GCTTTCTTCAAGACAGCAGTCCCGGAGGTTGGAAGCAAGGCTGAGTACCTGGGCTGGAGCGGAAC 1248  
Db 1149 GCTTTCTTCAAGACAGCAGTCCCGGAGGTTGGAAGCAAGGCTGAGTACCTGGGCTGGAGCGGAAC 1208  
Qy 1249 ATTCTCTTAAAGTCCGGAAGGAGATGACATGGCTCAAAATCTCAAGGTGACCCGG 1308  
Db 1209 ATTCTCTTAAAGTCCGGAAGGAGATGACATGGCTCAAGTCTGAGACTCTGAGAAATGACCCAG 1268  
Qy 1309 CTGAGGATACCTCTCTCGCCCTGGCAAGAGCAGCTTAAAGCTTCCGAAAGGCAATCTCA 1368  
Db 1269 TTGAAGATACCTCTCTCGCCCTGGCAAGAGCCTCAAGCTTCCGAAAGGATCTCTCA 1328  
Qy 1369 AGAAAGTCTCTCTCTGAGGAGGATGACAGGAGGACCTTCAAGAACTCAGACCGG 1428

Db 1329 AGAAAAAGGCTCTCCCTCATCGGGGAGGTACAGGAGGCCCTCAGGAATCAGACCAG 1388  
Qy 1429 TGCTGTATCTCAGGGCAGCTGTCTCCCTGTGTATCCCTGTCTCCCAAGGAAGGATCC 1488  
Db 1389 TGTCCAATACCCCAAGGCGCTGTCTCTCTATACCCCTGTCTCCCAAGGAAGGCAATC 1448  
Qy 1489 TTAAGAGTCTCAGACGCTGAATCTGTTTACTACTCTCTCAGAGCCAGGAGTCTG 1548  
Db 1449 TTAAGAGTCTCGGACGCTGAATCTGTTTACTACTCTCTCAGAGCCAGTGAATCTG 1508  
Qy 1549 GGGAACTCTTAGACCGCAGTGTGTGTGTAGTGGGACCCGCTGGAGCAGAAATCTC 1608  
Db 1509 GGGAACTCTTAGACCGCAGTGTGTGTGTAGTGGGACCCGCTGGAGCAGAAATCTC 1568  
Qy 1609 CACAGGCTCAGGGCTCTCTCCACCGCAAGGCAATTTCAAACTCAATGGCAAGTTCT 1668  
Db 1569 CACAAGCTTCAGGG---CGCTCATCGCAAGGCAATCTCAAACTCAATGGCAAGTTT 1625  
Qy 1669 CCGGCAAGCTTAGAGGCACTACCCCTAGCAGCTTTGGCTCCCTGGACCAACTGGCCT 1728  
Db 1626 CCGGCAAGCTTAGAGGCACTACCCCTAGCAGCTTTGGCTCCCTAGACCAACTGGCCT 1685  
Qy 1729 CTTCCATCTCGCAGCCCGGCCAGCGCCCTCAGGGCTGTGAGTGAAGCAGCATCC 1788  
Db 1686 CCGCTCATCTACAGCCCGGCCAGCGCTCCCTCGGAGCTGTGAGTGAAGCAGCATCC 1745  
Qy 1789 TGTCTCCGAGTCTTTGACCAATTTGCACTTTGGCTGAACTTTCCGAAACCCCACTGA 1848  
Db 1746 TGTCTCCGAGTCTTTGACCAATTTGCACTTTGGCTGAACTTTCCGAAACCCCACTGA 1805  
Qy 1849 GGGCTGTGTCTGTGGCAACCTCAGGGGCTTGGAGCAGCTCTCAGAGGCTGTA 1908  
Db 1806 GAGCTGTGTCTGTGGCAACCTCAGGGGCTTGGAGCAGCTCTCAGAGGCTGTA 1865  
Qy 1909 AGCGATGGTGGCAGGAATCTTTGGGGATAGCTGTCTTTCTCTGACAGACTGCGCAAGAGG 1968  
Db 1866 AACGATGGTGGCAGGAATCTTTGGGGATAGCTGTCTTTCTCTGACAGACTGCGCAAGAGG 1925  
Qy 1969 TGACTGAGGCTTACAGCAAGCCCTAGGAATCTGTCTCAAGCTCAGTGAAGGAGAGA 2028  
Db 1926 TGAAGCAGGCTTACAGCAAGCCCTAGGAATCTGTCTCAAGCTCAGTGAAGGAGAGA 1985  
Qy 2029 TGTGCTCCCTAGT--ATGGGTGAGTCTCTGAGAGGTTTGCAGAGAACTCTGGGTCCGAT 2087  
Db 1986 CAGTGCCCAAGTATGGGTGAGTCTCTTAGAGGGTTTGCAGAGAACTCTGGGT--AGATT 2044  
Qy 2088 CTTCCAGTGAATAGATACATCAAG---GGCTCTACGTCTGAGGCTCTGACTGAACCTGA 2143  
Db 2045 CCGCAGGTTGTAGAGTACATCAAGAACTCTCTCTCTGTCTTCTGAGCTGATGAACCTGG 2104  
Qy 2144 AGATGAGAGAAATCCATTTGATGTGGAAGGAATGGGAACCTTGTCTGCCCGAGTGTGA 2203  
Db 2105 AGGCTGAGAGAAATAGCAGAGATATGGAAGGACTGACCTACAGAGTCTGACTGCAAGATG 2164  
Qy 2204 TAGTGGGCT--GGCCTGAAGGTGCTCTCTCTTGTGCACTGAGTGTCAACCATGACA 2260  
Db 2165 TGAGCGCAGCAGAGACTGAAGTGGCTTCTCTCTTATGCT--TGAGTGTCTACCCATGGCA 2222  
Qy 2261 TTTCCACCCCTGTCTCTGGCTGCACC-----TTACATAAGTTTCTGTTTCCATCA 2312  
Db 2223 TCTCCC--CCCTGTCTCTGCCAGTGTGAGGTTATCCACATAAGTTCTCTGTTCCGATCG 2281  
Qy 2313 ACCACAGGTTAGAACCTGACTTCTCGGAGGTTAATGTAGTGAATCTGAGTGTGAGTGTGA 2372  
Db 2282 ACCACAGGTTAGAACCTGACTTCTCGGAGGTTAATGTAGTGAATCTGAGTGTGAGTGTGA 2341  
Qy 2373 GAGAGAAACAGCCTCTGTGTTTCCATCTCTGTGTCTGTGTCATCTCAAGACCTGGGAAGA 2432  
Db 2342 AGAGAGAAACAGCCTCTGTGTTTCCATCTCTGTGTGTGTCATCTCAAGACCTTAGAAGA 2401  
Qy 2433 CTGGACCGCTGTGTGACTTTCATCTCAAGGGAACAGATGCGCCCTGGACCCCATCTTAGA 2492  
Db 2402 CTC-AACTGCTGTTTCACTTTCATCTCAAGGGGAC-----2434







30-JUN-2000; 2000US-00608921.  
31-JUL-2000; 2000US-0221923P.  
25-SEP-2000; 2000US-0234922P.  
25-SEP-2000; 2000US-0235035P.  
08-NOV-2000; 2000US-0245669P.  
09-NOV-2000; 2000US-0071121P.  
14-NOV-2000; 2000US-0248325P.  
15-NOV-2000; 2000US-0248893P.  
22-DEC-2000; 2000US-0257511P.  
05-JAN-2001; 2001US-0260166P.  
28-FEB-2001; 2001US-0079703P.  
27-APR-2001; 2001US-00845044.  
20-JUL-2001; 2001US-00909743.  
31-JUL-2001; 2001US-00920346.  
13-AUG-2001; 2001US-00928531.  
14-AUG-2001; 2001US-00929218.  
15-AUG-2001; 2001US-0312539P.  
25-SEP-2001; 2001US-00963159.  
08-NOV-2001; 2001US-00080016.  
13-NOV-2001; 2001US-00012055.  
15-NOV-2001; 2001US-000303690.  
30-JAN-2002; 2002US-00060763.  
25-MAR-2002; 2002US-00105989.  
12-APR-2002; 2002US-00121911.  
12-AUG-2002; 2002US-00217168.  
22-OCT-2002; 2002US-00278036.  
02-JAN-2003; 2003US-00336489.  
03-JAN-2003; 2003US-00336153.  
  
(MILL-) MILLENNIUM PHARM INC.  
  
Kapeller-Libermann R, Hunter JJ, Meyers RE, Rudolph-Owen LA;  
Curtis RAJ, Olandt PJ, Teai F, Galvin KM, Chun M, Williamson MJ;  
Silos-Santiago I, Bandaru R;  
  
WPI: 2004-268788/25.  
P-PSDB; ADL14161.  
  
New 21910, 56634, 55053, 2504, 15977, 14760, 25501, 17903, 3700, 21529,  
26176, 26343, 56638, 18610, 33217, 21967, h1983, m1983, 38555 or 593  
nucleic acid molecules and proteins, useful for treating, e.g. cancer,  
heart failure and angina.  
  
Claim 1; SEQ ID NO 43; 139pp; English.  
  
The invention describes an isolated 21910, 56634, 55053, 2504, 15977,  
14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217,  
21967, h1983, m1983, 38555 or 593 nucleic acid molecule (I) comprising  
any one of 40 nucleotide sequences (I). The nucleic acid molecules and  
polypeptides are useful for diagnosing and treating a subject having a  
disorder, or a subject at risk of developing a disorder, which is  
associated with aberrant 21910, 56634, 55053, 2504, 15977, 14760, 25501,  
17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217, 21967, h1983,  
m1983, 38555 or 593 activity, such as cellular proliferative and/or  
differentiative disorders, brain disorders, platelet disorders, breast  
disorders, colon disorders, kidney (renal) disorders, lung disorders,  
ovarian disorders, prostate disorders, cervical disorders, spleen  
disorders, thymus disorders, thyroid disorders, testes disorders,  
hematopoietic disorders, pancreatic disorders, skeletal muscle disorders,  
skin (dermal) disorders, disorders associated with bone metabolism,  
immune, e.g. inflammatory disorders, cardiovascular disorders,  
endothelial cell disorders, liver disorders, viral diseases, pain  
disorders, metabolic disorders, neurological or central nervous system  
disorders, erythroid disorders, cancer, heart failure, hypertension,  
disorders (all claimed), e.g. cancer, heart failure, hypertension,  
angina, osteoarthritis, rheumatoid arthritis, multiple sclerosis, Crohn's  
disease, psoriasis, or asthma. The nucleic acid molecules and  
polypeptides are also useful as modulating agents in regulating a variety  
of cellular processes, e.g. cell proliferation, differentiation, growth and  
division. This sequence encodes a novel human protein of the invention.  
Note: The sequences given in the specification are also available in  
electronic format from  
ftp.segdata.uspto.gov/sequence.html?DocID=20040058355.

Db 1064 TCAACTGGGCTACGCCACCCAGTGGGAGAGCAGGAGCTCCGATGAGGGTGGGCACC 1123  
Qy 1102 CTAGTGTGTGATTTGGCCGGGCTCCATGGCGGACTGGTTAGCTCGCTCTCTCGGCCCCCC 1161  
Db 1124 CTGCGAGTGTACTGCGCGGCTCCATGGCTGACTGGCTCCGGCTCTCTCCCGCCCC 1183  
Qy 1162 TCCTGGAGATGGAGCAGGAGTGTGAGCTTCTTCAAGCAGCAGCTGCGCGGAGTGGAA 1221  
Db 1184 TCCTGGAGATGGGCGCAGGAGTGTGAGCTTCTTCAAGCAGCAGTGCACCTGGTGGGGAA 1243  
Qy 1222 GCACTGTACCTGGGCTGGAGCGGCAACATCTCTTAAAGATCCCGAAAGAGAGATGACA 1281  
Db 1244 GCACCACTCTGGGCTGGAGCGGCGAGCATTCGCTCAAGAAGTCCCGCAAGGAGATGACA 1303  
Qy 1282 TGGCTCAAAATCTGCAAGTGTACCGGCTGAGGATACCTCTTCTGCGCTGGCGAAGAGCA 1341  
Db 1304 TGGCCCAAGTCTCTCCAGTGTACACGGCTGATGACACTGCCCCATCGCCCTGGCAAGAGCA 1363  
Qy 1342 GCCTTAAGCTTCCGAAGGCAATCTCAAGAAAGTCTCTTACCTCTGTCAGGGGAGGTAC 1401  
Db 1364 ACCTCAAGTGTCCAAAGGCAATCTCAAGAAGAGTGTGAGCTCTGCAAGAGGGTAC 1423  
Qy 1402 AGGAGACCTTCAGAACTCAGACCGGTGCTGATCTTCCAGGCGAGCTGTCCCTGCTG 1461  
Db 1424 AGGAGACCTTCCGAGCTCAGGCCAATCCCTGCGAGCCAGGCGAGGCTGCC----- 1477  
Qy 1462 TATCCCTGCTCCCAAGGAGGCAATCTTAAAGATCTGACAGCGTGAATCTGGTACT 1521  
Db 1478 ---CCCTGCTCCCAAGAGGGCAATCTTCAAGAAGCCCCGACAGCGGAGTCTGGCTACT 1534  
Qy 1522 ACTCCTCTCCAGAGCCAGCGAGTCTGGGGAATCTTAGAGCCAGTGAATGTTCTGA 1581  
Db 1535 ACTCCTCTCCGAGCCAGTGAATCTGGGAGCTCTTGGAGCAGAGGCACTGTTTGTGA 1594  
Qy 1582 GTGGGACCCGCTGGAGCAGAAAGTCTCCAGGCTTCAGGCTCTCTCCACCGCAAG 1641  
Db 1595 GTGGGATCCCAAGGAGCAGAAAGTCTCCGAAAGTTCAGGGCTCTCTCCATCGCAAG 1654  
Qy 1642 GCATTCCTCAACTCAATGGCAAGTCTCCGCAAGCTTAGAAGCACTACCTTAGCA 1701  
Db 1655 GCATTCCTCAACTCAATGGCAAGTCTCCGCAAGCTTAGAAGCTCGCGGCCCCACCA 1714  
Qy 1702 CCTTTGGCTCCCTGGACCAACTGGCTCTCTCCATCTGAGCCGCGCCAGCGCCCT 1761  
Db 1715 CCTTGGCTCCCTGGATGAATCGCCCACTCGCCCTCGCCCGGCGCAGCGACCT 1774  
Qy 1762 CAGGGCTGTGAGTGGAGCAGCATCTCTGCTCCGAGTCTTTGACCAATGAGCTTC 1821  
Db 1775 CAGGGCTGTGAGCGAGGACAGCATCTCTGCTGAGTCTCTTACAGCTGAGCTTGC 1834  
Qy 1822 CTGAACCTCTTCCGAACCCCTGAGGGGCTGTGTCTGTGACAACTGAGGGGC 1881  
Db 1835 CTGAACGGCTCCAGAGCCCTCTGCGGGCTGTGTCTGTGGAACAACCTCAGCGGC 1894  
Qy 1882 TTGAGCAGCTCTCTCAGAAG-----GTCTGAAGCAGATGTGTGAGGAATCT 1929  
Db 1895 TTGAGGAGCCCTCTCAGAGGGCCCTGGAAGTCTCTGAGGCGCTGCGCGAGATCTT 1954  
Qy 1930 TGGGGATAGCTGCTTTCTGACAGACTGCCAAGAGGTGACTGACGCTACAGACAAG 1989  
Db 1955 TGGGGGACAGCTGCTTTCTGACAGACTGCCAGAGGTGACAGCGACCTACCGCAGG 2014  
Qy 1990 CCTTAGGAATCTGCTCAAGCTCAGCTGAGAGGGAGATGTGCTCTAGTATGGGGTAG 2049  
Db 2015 CATGAGGGTCTGCTCAAGCTCACCCTGAGTGAAGTAGGATGCCCCAG-CCCGGTGAG 2073  
Qy 2050 GCTCTGAGAGGGTTTGAGAGGAACCTTGGGTGGATTTCTCTCC 2092  
Db 2074 GCTCTCAGATGCAGCTGTTTGCACCCCGAGGGGAGATGCTCTC 2116

ID AAK94280 standard; cDNA; 3395 BP.  
XX  
AC AAK94280;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human full-length cDNA, SEQ ID NO: 2919.  
XX  
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.  
XX  
OS Homo sapiens.  
XX  
PN EPI130094-A2.  
XX  
PD 05-SEP-2001.  
XX  
PF 07-JUL-2000; 2000EP-00114089.  
XX  
PR 08-JUL-1999; 99JP-00194486.  
PR 11-JAN-2000; 2000JP-00118774.  
PR 02-MAY-2000; 2000JP-00183765.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Nishikawa T, Isogai T, Hayaashi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX  
DR WPI; 2001-524255/58.  
DR P-PSDB; AAM93360.  
XX  
PT 830 Primers useful for synthesizing full length cDNA clones and their use  
in genetic manipulation.  
XX  
PS Claim 8; SEQ ID NO 2918; 1380pp + Sequence Listing; English.  
XX  
CC The invention relates to primers for synthesising full length cDNA  
clones. 830 cDNA molecules encoding a human protein have been isolated  
and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have  
been determined. Primers for synthesising the full length cDNA are useful  
for clarifying the function of the protein encoded by the cDNA. The full  
length clones were obtained by construction of full length enriched cDNA  
libraries that were synthesised by the oligo-capping method. The primers  
enable the production of the full length cDNA easily without any special  
methods. The present sequence is a full length human cDNA of the  
invention. Note: The sequence data for this patent did not form part of  
the printed specification, but was obtained in CD-ROM format directly  
from EPO  
XX  
SQ Sequence 3395 BP; 722 A; 1044 C; 894 G; 735 T; 0 U; 0 Other;

Query Match 47.4%; Score 1375.8; DB 4; Length 3395;  
Best Local Similarity 81.7%; Pred. No. 4.2e-294;  
Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4;  
Qy 82 CTACTCTGCTCCGACCCCACTCCACCTCGGGTCCCGCACCACCTGGAGTGGTGGCTTAC 141  
Db 77 CTACTGATTCCTGCTCGGCCCTTGTCTACCTCTGCTCGCCATGGAGTGGTGGTTTCG 136  
Qy 142 TCAGCGCCGAGCCAGGCTCCCTCGGCTTCGCGCTTCGAGAGCGCCCGCGCGC 201  
Db 137 CGCGGCTCCGCGCCCACTCCCTCGCGCCGACAGCTA-----GCCCGCGCGC 184  
Qy 202 TGGCGGACGGGTCTCATAGTGCCTTAACTCTGATGAAGAGCAGGCGGTGAAGCGGC 261  
Db 185 TGGCGGAAGGGTGTATCAAGTCCGCCAAGCCCTTATGAAGAAGCAGGCGGTGAAGCGGC 244  
Qy 262 ACCATCACAACAACAACCTCGCGCACCGCTAGAGTTCCTGGAGAGCTGGGCAAGGGCA 321  
Db 245 ACCACCAAGCACAACCTCGCGCACCGCTACAGTTCCTGGAGACCTTGGGCAAGGGCA 304  
Qy 322 CTACGGGAAGGTGAAGAAGGCAACGAGAGCTCGGGGCGTCTGTGGCCATCAAGTCCA 391  
Db 305 CCTACGGGAAGGTGAAGAAGGCGCGGAGAGCTCGGGGCGCTGTGTGGCCATCAAGTCAA 364

Qy	382	TCAGGAAAGACAAAATCAAAGATGAGCAGGATCTGTCTGCACATACGAGGGAGATTGAGA	441
Db	365	TCCGGAAGGACAAAATCAAAGATGAGCAAGATCTGTATGCACATACGCGAGGGAGATTGAGA	424
Qy	442	TCATGTTCTTCACTCAACACCCACACATCATGTGCCATCCATGAAGTGTTTGAGAAATGACA	501
Db	425	TCATGTATCACTCAACACCCCTCACATCATTTGCCATCCATGAAGTGTTTGAGAACAGCA	484
Qy	502	GCAAGATTGTGATTGTTCATGGAGTATGCCAGCCGAGGCGATCTGTATGATTAATCATCAGTG	561
Db	485	GCAAGATCGTGATCGTCATGGAGTATGCCAGCGGGCGACCTTTATGACTACATCAGCG	544
Qy	562	AGCGGCCACGGCTGAGTGAGCGGAGCGCAGGCATTTCTTCGACAGATCGTGTCTGCCCC	621
Db	545	AGCGCGACGAGCTCAGTGAGCGCGAAGCTAGGCATTTCTTCGCGCAGATCGTCTCTGCCG	604
Qy	622	TGCACTACTGCCACACAGAACGGGATCGTTCACCGAGATCTCAAGCTCGGAAACATCCTTC	681
Db	605	TGCACTATTGCCATTCAGAACAGAGTTGTCAACCGAGATCTCAAGCTCGGAAACATCCTCT	664
Qy	682	TAGATGCCAATGGAAACATCAAGATTGCTGACTTTGGCCCTCTCAACCTGTACCACAAG	741
Db	665	TGATGCCAATGGGAATATCAAGATTGCTGACTTCGCGCTCTCCAACTCTACCATCAAG	724
Qy	742	GCAAGTTCTCCAGACGTTCTGTGGGAGCCCTCTCTACGCCTCGCTTGAGATAGTCAAG	801
Db	725	GCAAGTTCTTCGACAGCATTTCTGTGGAGCCCTCTATGCCTGCCACAGATTGTCAATG	784
Qy	802	GGAGCCCTATGTGGGCCAGAGGTGACAGCTGGTCTCTGGGGTTTCTCCTGTACATCC	861
Db	785	GGAGCCCTACACAGGCCAGAGGTGACAGCTGGTCCCTGGGTGTTCTCCTCTACATCC	844
Qy	862	TGTTGCATGGACCATGCGCTTTTGACGGGAGGATCATAAAACTCGTGAAGCAATCA	921
Db	845	TGTTGCATGGACCATGCGCTTTTGATGGGCAATGACCATAGATCCTAGTGAACAGATCA	904
Qy	922	GTAAACGGGGTTACCGTGAGCGGCCCAAGCCGTCCGATGCTGTGGCCCTGATCCGGTGGC	981
Db	905	GCAACGGGGCTACCGGGAGCACCTAAACCCCTCTGATGCTGTGGCCCTGATCCGGTGGC	964
Qy	982	TGTTAATGTTGAACCCACCGCTCGGGCCACACTGGAGGATGTAGCCAGTCAATTGTGTGG	1041
Db	965	TGTTTGTGTTGAACCCACCGCCCGGCCACCTCTGGAGATGTGGCCAGTCACTGTGTGG	1024
Qy	1042	TCAACTGGGGTTACACCAACGGAGTGGGGAAACAGGAAGCCCTCGCTGAGGGTGGGCACC	1101
Db	1025	TCAACTGGGGCTACGCCACCGAGTGGGAGACAGGAGGCTCCCGCATGAGGGTGGGCACC	1084
Qy	1102	CTAGTGTGTGACTTTTGGCCGGGCGCTCCAATGGCGGACTGGTTACGTCTGCTCGCGCCCC	1161
Db	1085	CTGSCAGTGACTCTGCCCCGCGCTCCATGCGTGACTGCTCGCGGTCTCTCCCGCCCC	1144
Qy	1162	TCTTGGAGATGGAGCCAAAGGTGTGACGTTCTTCAAGACAGCAGTGCCCGGAGGTGGA	1221
Db	1145	TCTTGGAGAATGGGGCCAAAGGTGTGACGTTCTTCAAGACAGCATGCACTGGTGGGGAA	1204
Qy	1222	GCACGTACTGGGCTGGAGGGCAACATTTCTTTAAGAGTCCCGAAAGGAGAAATGACA	1281
Db	1205	GCACCAACCTCTGGGCTGGAGCGCCAGCATTTGCTTCAAGAGTCCCGCAAGGAGAAATGACA	1264
Qy	1282	TGGCTCAAAATCTGCAAGGTGACCCCGCTCAGGATACCTTTCTCGCCCTGGCAAGAGCA	1341
Db	1265	TGGGCCAGTCTCTCCACAGTGACAGGCTGATGACATGCCCCATGCGCTTGGCAGAGCA	1324
Qy	1342	GCCTTAAGCTTCGAAAGGCAATTTCTCAAGAAAAAGTCCCTTAAGTCTGCTCAGGGAGGTAC	1401
Db	1325	ACCTCAAGCTGCCAAAGGGCAATTTCTCAAGAAAGAGGTGTGAGCCTCTGCAAGAGGGGTAC	1384
Qy	1402	AGGAGGACCTTCAGGAACACTCAGACGGGTGCTGATATCCAGGGCAGCCTGTCCCTGCTG	1461
Db	1385	AGGAGGACCTTCGGAGCTCAGGCCAATTCCTTCGAGCCACAGGGCAGGCTGCC-----	1438

QY	1462	TATCCCTGCTCCCAAGGAAGGCATCTTTAAGAAAGTCTCGACAGCGTGAATCTGGGTACT	1521
Db	1439	--CCCTGCTCCCCAAGAAGGGCAATTCAGAAGACCCCGACAGCGGAGTCTGGCTACT	1495
QY	1522	ACTCCTCTCCAGAGCCCGAGGCTCTGGGGAACTCTTAGAGCGCCAGTGATGTGTTCTGA	1581
Db	1496	ACTCCTCTCCGAGCCCGAGTGAATCTGGGAGGCTCTTGGACGCGAGGCGACGTGTTGTGA	1555
QY	1582	GTGGGGACCCCGTGGAGCAGAAGTCTCCACAGGCTTCAGGGCTCTCCTCCACCGCAAGG	1641
Db	1556	GTGGGGATCCCAAGGAGCAGAAGGCTTCGCAAGCTTCAGGGCTCTCCTCCATCGCAAG	1615
QY	1642	GCATTCCTAAACTCAATGGCAAGTTCCTCCCGCACAGCCTTAGAAGGCATACCCCTAGCA	1701
Db	1616	GCATCCTCAAACTCAATGGCAAGTTCCTCCACAGACAGCCTTGGAGCTCGCGGCCCAACA	1675
QY	1702	CTTTTGGCTCCCTGGAGCAACTGGCCCTCTCCCATCTCGACGCCGCGCCAGCGGCCCT	1761
Db	1676	CTTTCGGCTCCCTGGATGAACCTGCCCCACCTGCCCCCTGGCCCCGGGCGAGCGACCT	1735
QY	1762	CAGGGCTGTGAGTGAGGACAGCATCCTGCTCCCGAGTCTTTTGACCAATTGGACTTGC	1821
Db	1736	CAGGGCTGTGAGCGAGGACAGCATCCTGCTCTGAGTCTTTTGACCAGCTGGACTTGC	1795
QY	1822	CTGAACGTCTTCCGAAACCCCACTGAGGGCTGTGTCTGTGGACAACTGAGGGGGC	1881
Db	1796	CTGAACGGCTCCCAAGAGCCCCCACTGCGGGCTGTGTCTGTGGACAACTCACGGGGC	1855
QY	1982	TTGAGCAGCTCCTCTCAGAG-----GTCTGAAGCGATGTGTGGCAGGAATCCT	1929
Db	1856	TTGAGGAGCCCCCTCAGAGGGCCCTTGAAGCTGCCTGAGGGCGCTGGCGCAGGATCCTT	1915
QY	1930	TGGGGGATAGCTGCTTTTCTCTGACACACTGCCAAGAGTGACTGCGAGCCTACAGACAAG	1989
Db	1916	TGGGGGACAGCTGCTTTTCTCTGACACACTGCCAAGAGTGACTGCGAGCCTACCGACAG	1975
QY	1990	CCCTAGGAATCTGCTCAAAAGCTCAGCTGAGGAAGGGAGATGGTGCCCTAGTATGGGGTAG	2049
Db	1976	CACCTGAGGCTGTCTCAAAGCTACCTGAGTGGAGTAGGCATTGCCCCAG-CCCGGTGAG	2034
QY	2050	GCTCTGAGAGGTTTGACAGAGGAACCTTGGTGGATTCCTCC	2092
Db	2035	GCTCTGAGTGCAGCTGGTGTGCAACCCCGAGGGGAGATGCTTC	2077
RESULT 6			
ADL30885			
ID	ADL30885 standard; cDNA; 3395 BP.		
XX	AC ADL30885;		
XX	AC		
DT	20-MAY-2004 (first entry)		
XX	Full length human cDNA clone SeqID 2918.		
DE	human; medicine; signal transduction; glycoprotein; transcription;		
KW	oligo-capping method; ss; gene.		
OS	Homo sapiens.		
FN	EP1396543-A2.		
PD	10-MAR-2004.		
XX	07-JUL-2000; 2003EP-00025638.		
PR	08-JUL-1999; 99JP-00194486.		
PR	11-JAN-2000; 2000JP-00118774.		
PR	02-MAY-2000; 2000JP-00183865.		
PR	07-JUL-2000; 2000EP-00114089.		
XX	(REAS-) RES ASSOC BIOTECHNOLOGY.		
FA			
XX			

PI	Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;	
PI	Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;	
XX	WPI; 2004-204755/20.	
DR	P-PSDB; ADL30886.	
XX		
PT	New oligonucleotide primers (830 cDNAs) useful for synthesizing full	
PT	length human cDNAs.	
XX		
XX	Example 1; SEQ ID NO 2918; 1340pp; English.	
CC	This invention relates to a novel primers useful for synthesizing full	
CC	length cDNA molecules that encode human proteins. Specifically, it refers	
CC	to secretory or membrane proteins that are potential therapeutic agents/	
CC	target molecules in the field of medicine, and in particular genes	
CC	encoding proteins that are associated with signal transduction,	
CC	glycoproteins and transcription. The present invention describes a method	
CC	for efficiently cloning a full length human cDNA from both the 5' and 3'	
CC	ends using the oligo-capping method. This polynucleotide sequence is a	
CC	full length human cDNA clone of the invention.	
XX		
QQ	Sequence 3395 BP; 722 A; 1044 C; 894 G; 735 T; 0 U; 0 Other;	
	Query Match 47.4%; Score 1375.8; DB 12; Length 3395;	
	Best Local Similarity 81.7%; Pred. No. 4.2e-294;	
	Matches 1652; Conservative 0; Mismatches 337; Indels 34; Gaps 4	
QY	82 CTACTGTCGCCGACCCACTCCACTCGCGGTCCCGCACCATGAGTCGTTGGCCTTAC 141	
DB		
DB	77 CTACTGATTCCTCTCGCGCCCTTGCTCACTCTGCTCGCCATGGAGTCGCTGGTTTCG 136	
QY	142 TCCAGCGCCGAGCGCAGGCTCCCTCGGCCTCGCCCTGGCCTCGAGAGCGCCGCGCGC 201	
DB		
DB	137 CGCGCGCTCGGCGCCCACTCCTTCGGCGGAGACTA-----GCCCGCGCGC 184	
QY	202 TGGCGGACGGGCTCATCAAGTCGCTAAACCTCTGATGAAGAAGCAGCGGTGAAGCGCG 261	
DB		
DB	185 TGGCGGAAGGGCTGATCAAGTCGCCCAAGCCCTTAATGAAGAAGCAGCGGTGAAGCGCG 244	
QY	262 ACCATCAACAACACAACTCGGGGCACCGCTACGAGTTCCTGGAGACGCTGGCGCAGGCA 321	
DB		
DB	245 ACCACCAAGCACAACTCGGGCACCGCTACGAGTTCCTGGAGACCTTGGGCAAGGCA 304	
QY	322 CCTACGGGAAGGTGAAGAAGCAGCAGAGAGCTCGGGCGCTCTGGTGCCATCAAGTCCA 381	
DB		
DB	305 CCTACGGGAAGGTGAAGAAGCGCGGGAGAGCTCGGGCGCTCTGGTGCCATCAAGTCAA 364	
QY	382 TCAGGAAGACAAAATCAAGATGAGCAGGATCTGCTGCACATACGAGGGAGATTGAGA 441	
DB		
DB	365 TCGGGAAGGACAAAATCAAGATGAGCAAGATCTGATGCAATACGGAGGAGATTGAGA 424	
QY	442 TCATGTCCTTCACTCAACACACCCCAATCATTTGCCATCCATGAAGTGTTCAGAAATAGCA 501	
DB		
DB	425 TCATGTCATCACTCAACACCCCTCACATCATTTGCCATCCATGAAGTGTTCGAGAACGCA 484	
QY	502 GCAAGATTGTGATTGTCTATGAGTATGCGACCGAGGCGATCTCTGTATGATTACATCATGTG 561	
DB		
DB	485 GCAAGATCGTATCGTCTATGAGTATGCGACCGGGCGACCTTTATGACTACATCAGCG 544	
QY	562 AGCGGCCACGCTCAGTGTAGCGGGACGCGCAGGCATTTCTCCGACAGATCTGTGTCTGCC 621	
DB		
DB	545 AGCGGCAGCAGCTCAGTGTAGCGGAGCTAGGCATTTCTTCGGCAGATCGTCTCTGCCG 604	
QY	622 TGCACTACTGCCACAGAAACGGGATCGTTTCAACGAGATCTCAAGCTGGAACATCTCTTC 681	
DB		
DB	605 TGCACTATTGCCATTCAGAACACAGATTGTGCCACCGAGATCTCAAGCTGAGAACATCTCT 664	
QY	682 TAGATGCCAATGGAAACATCAAGATTCTGTGACTTTGGCCTCTCCAACTGTACCAACAAG 741	
DB		
DB	665 TGAATGCCAATGGGAATATCAAGATTCTGTGACTTCTGGCGCTCTCCAACCTCTACCATCAAG 724	
QY	742 GCAAGTTCCTCCAGACGCTTCTGTGGGAGCCCTCTCTACGCTCTCGCTGAGATAGTCAACG 801	

QY 1882 TTGAGCAGCCTCCCTCAGAG-----GTCTGAAGCGATGGTGGCAGGAATCCT 1929  
 DB 1856 TTGAGGAGCCCCCTCAGAGGGCCCTGGAAGCTGCCTGAGGCGCTGGCGCAGGATCCTT 1915  
 QY 1930 TGGGGGATAGCTGTTTCTGTGACAGACTGCCAAGAGGTGACTGCGAGCCTACAGACAAG 1989  
 DB 1916 TGGGGGACAGCTGCTTTTCTGACAGACTGCCAGGAGGTGACAGCGACCTTACCAGCAGG 1975  
 QY 1990 CCCTAGGAATCTGCTCAAAAGCTCAGCTGAGGAAGGAGATGGTCCCTAGTATGGGCTAG 2049  
 DB 1976 CACTAGGGTCTGCTCAAGCTCACTGAGTGGAGTAGGCATTCGCCAG-CCCGGTGAG 2034  
 QY 2050 GCTCTGAGAGGGTTTGACAGAGGACCTGGGTGCGATTCTCTCC 2092  
 DB 2035 GCTCTCAGATGACGTGGTGGACCCCGAGGGGAGATGCGCTTC 2077

## RESULT 7

ABX71420

ID ABX71420 standard; cDNA; 3443 BP.

XX AC ABX71420;

XX DT 14-APR-2003 (first entry)

XX Human cell cycle-associated cDNA from clone DKFZphtes3\_7j3.

XX Human; gene; gene therapy; vaccine; disease treatment; detection; ss.

XX OS Homo sapiens.

XX PN WO200112659-A2.

XX PD 22-FEB-2001.

XX PF 18-AUG-2000; 2000WO-IB001496.

XX PR 18-AUG-1999; 99US-0149499P.

XX PR 28-SEP-1999; 99US-0156503P.

XX PA (GEHU-) GERMAN HUMAN GENOME PROJECT.

XX PI Wiemann S;

XX DR WPI; 2001-327840/34.

XX DR P-PSDB; ABUS3319.

XX PT Nucleic acids having the sequences of clones isolated from libraries of different human tissues, useful in recombinant DNA methodologies.

XX PS Claim 1; Page 942-943; 1095pp; English.

XX CC This invention describes novel polynucleotides and polypeptides isolated from human cDNA libraries which can be used for gene therapy or in vaccines. The polynucleotides of the invention and antibodies encoded by them may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate polypeptide expression. The products of the invention may also be used to identify modulators of expression and activity and to down regulate expression and activity. The antibodies of the invention may also be used as diagnostic agents for detecting the presence of polypeptides in samples. This sequence encodes a polypeptide described in the disclosure of the invention.

XX SQ Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;

Query Match 47.4%; Score 1374.2; DB 5; Length 3443;  
 Best Local Similarity 81.6%; Pred. No. 9.5e-294;  
 Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;

QY 82 CTACTGCTGCCGACCCACCTCCACCTCGCGGTCCCGCACCATGGAGTCGGTCCCTTAC 141  
 DB 87 CTACTGATTCCTCGCGCCCTTGCTCACCTCTGCTCGCCATGGAGTCGCTGGTTTCG 146

QY 142 TCAGCGCCCGCAGCAGGCTCCCTCGGCCTCGCCCTCGGAGCGCCCGCGGC 201  
 DB 147 CGCGCGCTCGGCGCCCACTCCCTCGGCGCAGAGCTA-----GCCCGCGGC 194  
 QY 202 TGGCGGACGGGCTCATCAAGTCGCTAAACCTCTGATGAAGAGCAGCGGTGAAGCGGC 261  
 DB 195 TGGCGGAAGGGCTGATCAAGTCGCGCAAGCCCTTAATGAAGAGCAGCGGTGAAGCGGC 254  
 QY 262 ACCATCAAAACACACACCTCGGCGCAGCGCTACGAGTTCTTGGAGACGCTGGGCGAGGCA 321  
 DB 255 ACCACCAACAGCAACACCTCGCGCACCGCTACGAGTTCTTGGAGACCTGGGCGAAGGCA 314  
 QY 322 CCTACGGGAAGGCTGAAGAGCGCACGAGAGAGCTCGGGCGCTCGGTGGCCCATCAAGTCCA 381  
 DB 315 CCTACGGGAAGGCTGAAGAGCGCGGGAGAGCTCGGGCGCTCGGTGGCCCATCAAGTCAA 374  
 QY 382 TCAGGAAGACAAAAATCAAAAGATGAGCAGGATCTGTGCAATACGAGGAGATTGAGA 441  
 DB 375 TCCGGAAGGACAAATCAAAAGATGAGCAAGATCTGATGCACATACGAGGAGGATTGAGA 434  
 QY 442 TCATGCTTCTCACTCAACCAACCCCAACATCATTCGCCATCCATGAAGTGTGAGNATAGCA 501  
 DB 435 TCATGTCATCACTCAACCAACCCCTCACATCATTCGCCATCCATGAAGTGTGAGNATAGCA 494  
 QY 502 GCAAGATTGTGATTGTTCATGGAGTATGCCAGCGGAGCGGATCTGTATGATTACATCAGTG 561  
 DB 495 GCAAGATCGTATCGTATGAGATATGCCAGCGCGGGGACCTTTATGACTACATCAGCG 554  
 QY 562 AGCGGCCACGGCTGAGTGAGCGGAGCGCCAGGCAATTTCTTCCAGCAGATCGTGTCTGCC 621  
 DB 555 AGCGGCAGCAGCTCAGTGAGCGGAACTAGGCAATTTCTTCCGCGAGATCGTCTCTGCCG 614  
 QY 622 TGCACTACTGCCACAGAGCGGATCGTTTCCAGAGATCTCAAGCTGGAAAACATCTCTTC 681  
 DB 615 TGCACTATTGCCATCAGAACAGAGTTGTCCACCGAGATCTCAAGCTGGAGAATCTCTCT 674  
 QY 682 TAGATGCCAATGGAACATCAAGATTGTCGACTTTGGCCCTCTCCAACCTGTACCACAAAG 741  
 DB 675 TGATGCCAATGGGATATCAAGATTGTCGACTTCGGCCCTCTCAACCTGTACATCAAG 734  
 QY 742 GCAAGTTCCTCCAGACGTTCTGTGGGAGCGCTCTCTACGCTCTCGCTCGATAGATAGTCAACG 801  
 DB 735 GCAAGTTCCTGCAGACATTCGTGGGAGCGCCCTCTATGCTCTGCCAGAGATTGTCAATG 794  
 QY 802 GGAAGCCCTATGTGGGCGCCAGAGTGGACAGCTGGTCTCTGGGCGGTCTCTCTGTACATCC 861  
 DB 795 GGAAGCCCTTACACAGGCCCCAGAGTGGACAGCTGGTCCCTGGGTGTCTCTCTTACATCC 854  
 QY 862 TGGTGCATGSCACCATGCCCTTTGACGGGCGAGGATCATAAACACATGGTGAAGCAAAATCA 921  
 DB 855 TGGTGCATGGCACCCTATGCCCTTTGATGGGCGATGACCATAGATCTTAGTGAACAGATCA 914  
 QY 922 GTAACGGGGCTTACCGTGAAGCGCCCAAGCGCTCGATGCTGTGGCTGATCCGGTGGC 981  
 DB 915 GCAACGGGGCTTACCGGGAGGCCACCTAAACCTCTGATGCTGTGSCCTGATCCGGTGGC 974  
 QY 982 TGTTAATGTTGAACCCCAACCGCTCGGGCGCACACTGAGAGNATGTAGCCAGTCAATTTGGTGG 1041  
 DB 975 TGTGTATGTTGAACCCCAACCGCTCGGGCGCACCTCGGAGGATGTGGCCAGTCACTGTGTGG 1034  
 QY 1042 TCAACTGGGGTTACACCAACCGGAGTCGGGGAAACAGAGACCCCTGCGTGAAGGGTGGGCACC 1101  
 DB 1035 TCAACTGGGGCTACGCCACCCGAGTGGGAGACGAGGAGGCTCCGCATAGGGTGGGCACC 1094  
 QY 1102 CTAGTGGTGAATTTGGCCGGGCGCTCCATGCGGAGCTGGTTACGTCGCTCTCTCGCGCCCC 1161  
 DB 1095 CTGGCAGTGACTCTGCGCGCGCTCCATGAGTGGCTCCGCGCTTCTCTCCCGCCCC 1154  
 QY 1162 TCTGAGAGATGAGGCGCAAGGTGTGAGCTTCTTCAAGCGACGAGTGTCCGGGAGGTGGAA 1221  
 DB 1155 TCTGAGAGATGCGGGCGCAAGGTGTGAGCTTCTTCAAGCGACATGCACTGTGTGGGGGAA 1214









PA (HAFE/) HAERLACH T.  
PA (SCHO/) SCHOCH C.  
PA (KERN/) KERN W.  
XX  
PI Haerlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;  
PI Bils R, Brors B, Mergenthaler S;  
XX  
XX WPI; 2003-505037/47.  
XX  
XX Determining the subtype of leukemia cells and whether a patient sample  
PT contains leukemia cells or other cells, useful for treating leukemia,  
PT comprises determining the expression profile of a group of markers in a  
PT patient sample.  
XX  
XX Disclosure; SEQ ID NO 2508; 2938pp; English.  
XX  
XX The present invention relates to a method (M1) for determining the  
CC subtype of leukemia cells and whether a patient sample contains  
CC leukemia cells. The method comprises determining the expression profile  
CC of a group of markers in a patient sample. The method is useful for  
CC determining the presence of leukemia cells, its types or subtypes, and  
CC for the preparation of a medicament for treating leukaemia.  
XX  
XX Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;  
SQ  
Query Match 47.4%; Score 1374.2; DB 10; Length 3443;  
Best Local Similarity 81.6%; Pred. No. 9.5e-294; Indels 34; Gaps 4;  
Matches 1651; Conservative 0; Mismatches 338;  
82 C T A C T G C T G C C G A C C A C C A C T C C A C C T C G C G G T C C C C G C A C C A T G G A G T C G G T G C C T T A C 141  
87 C T A C T G A T T C C C C T G C C G C C T T G C T C A C C T C C T G C T G C C A T G A G T C G C T G G T T T C G 146  
142 T C A G C C C C G A G C C A G G T C C C T G C G C C T C C G C C T G C C C G C A C C A T G G A G T C G G T G C C T T A C 201  
147 C G C G G C G C T C G G C C C A C C T C C T C G C G C G C A G A G T A - - - - - G C C C G G C C G C 194  
202 T G G C G A C G G G C T C A T C A A G T C G C C T A A A C T C T G A T G A A G A C A G C G C G T G A A G C G C 261  
195 T G G C G A G G G C T G A T C A G T C G C C A A G C C C T A A T A G A A G A C A G C G C G T G A A G C G C 254  
262 A C C A T C A A A C A A C A C C T G G G C A C C G T A C A G T T C C T G G A G A C C T G G C A A G G G C A 321  
255 A C C A C A C A A G A C A A C C T G C G G C A C C G T A C A G T T C C T G G A G A C C C T G G G C A A A G G C A 314  
322 C C T A C G G A A G T G A A G A A G C G C A G A G C T C G G G C G T C T G G T G C C A T C A A G T C C A 381  
315 C C T A C G G A A G T G A A G A A G C G C G G A G A G C T C G G G C G C C T G T G G C C A T C A A G T C A A 374  
382 T C A G G A A G A C A A A T C A A A G A T G A G A G A T C T G T G C A C A T A C G G A G G A G A T T C G A 441  
375 T C G G A A G A C A A A T C A A A G A T G A G A G A C A G A T C T G A T G C A C A T A C G G A G G A G A T T G A 434  
442 T C A T G C T T C A C T C A A C C A C C C C A C A T A T T G C C A T C C A T G A A G T T T T G A G A A T A G C A 501  
435 T C A T G T C A T C A C T C A A C C C C A C A C C T C A C A T A T T G C A T C C A T G A A G T G T T T G A G A A C A G C A 494  
502 G C A A G A T T G A T T G T C A T G A G A T G C C A G C C G A G G A T C T G A T G A T T A C A T C A G T G 561  
495 G C A A G A T C G T G A T G C T A T G A G A T A T C C A G C G G G C G A C C T T A T G A C T A C A T C A G C G 554  
562 A G C G G C C A C G G C T G A T G A G C G G A C C C A G G A T T T T T C C G A C A G A T C G T G T G C C C 621  
555 A G C G G C A G A C T C A G T G A G G C G A G T A G G A T T T T T C C G G C A G A T C G T C T G C C G 614  
622 T G C A T C T C C C A C C A A G A C G G A T C G T T C A C G A G A T C T C A A G T G G A A A A C A T C C T T C 681  
615 T G C A C T A T T G C C A T C A A C A G A G T T G T C C A C C A G A T C T C A A G C T G A G A A C A T C C T C T 674  
682 T A G A T G C C A T G A A A C A T C A A G A T T G C A C T T G G C C T C C A C C T G T A C C A A A G 741  
675 T G A T G C C A A T G G A T A T C A A G A T T G C A C T T G C G C C T C C A C C T C T A C C A T C A C A A G 734

742 G C A A G T T C C T C C A G A G T T C T G T G G A G C C C T C T C T A C G C C T C G C C T G A G A T A G T C A A C G 801  
735 G C A A G T T C C T G C A G A C A T T C T G T G G A G C C C C T C T A T G C C T C G C C A G A T T G T C A A T G 794  
802 G G A A G C C C T A T G T G G G C C C A G A G T G G A C A G C T G T C T T G G G C G T T C T C C T G T A C A T C C 861  
795 G G A A G C C C T A C A C A G G C C C A G A G T G G A C A G C T G T C C C T G G G T G T T C T C C T C T A C A T C C 854  
862 T G T G A T G G C A C A T G C C C T T T G A C G G G C A G A T A T A A A C A C A T G T G T A A G A A A T C A 921  
855 T G T G A T G G C A C A T G C C C T T T G A T G G G C A T A C C A T A G A T C C T A G T G A A A C A G A T C A 914  
922 G T A A C G G G G C T T A C C G T G A G C C C C A A G C C G T C C A T G C C T G T G C C T A C C G G T G G C 981  
915 G C A A C G G G G C T A C C G G G A C C A C C T A A A C C C T T G A T G C C T G T G G C C T A T C C G T G G C 974  
982 T G T T A A T G T G A A C C C C A C C C G T C G G G C C A C A C T G G A G A T G T A G C A G T C A T T G T G G G 1041  
975 T G T T A A T G T G A A C C C C A C C C G C C G G C C A C C C T G G A G A T G T G G C C A G T C A C T G T G G G 1034  
1042 T C A A C T G G G G T T A C A C A C C G G A G T C G G G A A C A G A A G C C C T G C T G A G G T G G G C A C C 1101  
1035 T C A A C T G G G G C T A C G C C A C C C G A G T G G G A G A G A G G C T C C G A T A G G G T G G G C A C C 1094  
1102 C T A G T G T G A C T T T G C C G G C C T C C A T G G C G A C T G T T A C G T C G C T C C T C G C G C C C C 1161  
1095 C T G C A G T G A C T C T G C C G G C C T C C A T G C T G A C T G G C T C C G G C G T T C T C C C G C C C C 1154  
1162 T C C T G G A A T G G A G C C A A G G T G T G A G C T T C T T C A A G C A G C A C G T C C G G G A G G T G G A A 1221  
1155 T C C T G G A A T G G G C C A A G G T G T G A G C T T C T T C A A G C A G C A T G C A C C T G T G G G G A A 1214  
1222 G C A C T G A C T G G G C T G G A G C G C A C A T T C T T A A G A A G T C C C A A A G A G A A T G A C A 1281  
1215 G C A C C A C C C T G G C C T G G A G C G C A G C A T T C G C T C A A G A A G T C C C A A G A G A A T G A C A 1274  
1282 T G C C T C A A A T C T G A A G T G A C C C G G C T G A G A T A C C T T C T C C C C T G G C A A G A C A 1341  
1275 T G G C C A G T C T C C A C A G T G A C A C G G C T G A T G A C A C T G C C C A T C G C C T G G C A A G A C A 1334  
1342 G C T T A A G C T T C C A A A G G A T T C T C A A G A A A A G T C C T T A C C T C G T C A G G G A G G T A C 1401  
1335 A C C T A A G C T G C C A A A G G C A T T C T C A A G A A A G G T G T C A G C C T C T G C A A A A G G G T A C 1394  
1402 A G A G A C C C T C A G A A C T C A G A C C G T G C C T G A T A C T C C A G G C A G C C T G T C C C T G C T G 1461  
1395 A G A G A C C C T C G G A G C T C A G C C C A A T C C C T G C A G C C A G G C A G G C T G C C C - - - - - 1448  
1462 T A T C C C T G C T C C C A A G A A A G G C A T C C T T A A G A A G T C T G A C A G C G T G A A T C T G G T T A C T 1521  
1449 - - - C G C T G C T C C C A A G A G G C A T T C T C A G A A G C C C G A C A G C G G A G T C T G G C T A C T 1505  
1522 A C T C C T C T C A G A C C C A G C G A G T C T G G G A A C T T T A G A C G C C A G T A G T G T T T G T G A 1581  
1506 A C T C C T C T C C G A C C C A G T G A A T C T G G G A G A G T C T T T G A C G C A G C A G C G T G T T T G T G A 1565  
1582 G T G G G A C C C C T G G A G C A G A G T C C C A C A G C C T T C A G G C T C C T C C C C A C C G A A G 1641  
1566 G T G G G A T C C C A A G A G A G A A G C C T C C C A G C A G C T T C A G G C T G C T C C A T C A T C G A A A G 1625  
1642 G C A T T C T C A A A C T C A A T G G C A A G T T C C C G C A C A G C C T T A G A A G G C A C T A C C C C T A G C A 1701  
1626 G C A T C T C A A C T C A A T G G C A A G T T C T C C A G A C A G C C T T G A G C T C G C G G C C C C A C C A 1685  
1702 C T T T G G C T C C C T G A C C A A C T G G C C T C C T C C A T C C T G A G C C C G G C C A G C C G C C C T 1761  
1686 C T T C G G C T C C C T G G A T G A A C T C G C C C A C C T C G C C C C C T G G C C C G G C C A G C C G A C C C T 1745  
1762 C A G G G C T G T G A G T A G A C A G A C A T C T G C T C C T C C G A G T C C T T T G A C C A A T T G G A C T T G C 1821  
1746 C A G G G C T G T G A G C G A G A C A G A C A T C C T G T C C T C T G A G T C C T T T G A C C A G C T G G A C T T G C 1805  
1822 C T G A A C G T C T C C C G A A A C C C A C C A C T A G G G G C T G T G T C T G T G G A C A A C C T G A G G G G C 1881

Db 1806 CTGAACGGCTCCAGAGCCGCCACTGCGGGGCTGTGTCTGTGGACAACTCAGGGG 1865  
Qy 1882 TTGAGCAGCTCCCTCAGAG-----GTCTGAAGCGATGTGGCAGGATCCT 1929  
Db 1866 TTGAGGAGCCCCCTCAGAGGGCCCTGGAAGCTGCCCTGAGGCGCTGGCGGAGATCCTT 1925  
Qy 1930 TGGGGGATAGCTGTTTTTCTGACAGACTGCCAAGAGGTGACTGACAGCCTACAGACAAG 1989  
Db 1926 TGGGGGACAGCTGTTTTTCTGACAGACTGCCAAGAGGTGACTGACAGCCTACAGACAAG 1985  
Qy 1990 CCTAGGAATCTGCTAAAGCTCAGCTGAGAGAGGAGATGGTGGCCCTAGTATGGGGTAG 2049  
Db 1986 CACTGAGGGTCTGCTCAAGCTCAGCTGAGTGGAGTAGGCAATGGCCCCAG-CCGGGTGAG 2044  
Qy 2050 GCTCTGAGAGGGTTTGAGAGGAACTTGGGTGCGATTCTCTCC 2092  
Db 2045 GCTCTGAGTGCAGCTGGTTGCACCCCGAGGGGAGATGCCTTC 2087

RESULT 10  
AD020171  
ID AD020171 standard; cDNA; 3443 BP.  
XX AC AD020171;  
XX XX  
DT 12-AUG-2004 (first entry)  
XX Human PRO polynucleotide #540.  
XX  
KW Human; PRO; gene; ss; immune related disorder;  
KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;  
KW juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;  
KW vasculitis; sarcoidosis; autoimmune haemolytic anaemia;  
KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;  
KW renal disease; demyelinating disease; central nervous system;  
KW peripheral nervous system; demyelinating polyneuropathy;  
KW Guillain-Barre syndrome;  
KW chronic inflammatory demyelinating polyneuropathy.  
XX  
OS Homo sapiens.  
XX  
PN WO2004043361-A2.  
XX  
PD 27-MAY-2004.  
XX  
PF 06-NOV-2003; 2003WO-US035268.  
XX  
PR 08-NOV-2002; 2002US-0425235P.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;  
PI Wood WI, Wu TD;  
XX  
XX WPI: 2004-420067/39.  
XX P-PSDB; AD020172.  
XX  
PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for  
PT treating an immune related disorder such as systemic lupus erythematosus,  
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or  
PT spondyloarthritis.  
XX  
PS Claim 1; SEQ ID NO 1136; 1731pp; English.  
XX  
CC The invention relates to human PRO polypeptides and the polynucleotides  
CC encoding them. The polypeptides and polynucleotides are useful for  
CC treating and diagnosing immune related disorders in mammals. The immune  
CC related disorders include systemic lupus erythematosus, rheumatoid  
CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic  
CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune  
CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes  
CC mellitus, immune-mediated renal disease, demyelinating diseases of the

CC central or peripheral nervous system, demyelinating polyneuropathy,  
CC Guillain-Barre syndrome and chronic inflammatory demyelinating  
CC polyneuropathy. This sequence represents a human PRO polynucleotide of  
CC the invention.  
XX  
SQ Sequence 3443 BP; 763 A; 1043 C; 896 G; 741 T; 0 U; 0 Other;  
Query Match 47.4%; Score 1374.2; DB 12; Length 3443;  
Best Local Similarity 81.6%; Pred. No. 9.5e-294;  
Matches 1651; Conservative 0; Mismatches 338; Indels 34; Gaps 4;  
Qy 82 CTACTGCTGCCGACCCACCTCCACCTCGCGGTCCCCGACCATGAGATCGGTGGCCTTAC 141  
Db 87 CTACTGATTCCCTGCGCCCTTGCTCACCCTCTGCTCGCATGAGATCGCTGTTTCG 146  
Qy 142 TCCAGCGCCGAGCAGGCTCCCTCGGCTCCGCTCGGCTCGGAGAGCGCCGCGCGC 201  
Db 147 CGCGGCGCTCCGGCCACCTCCCTCGGCGCAGAGCTA-----GCCGCGCGC 194  
Qy 202 TGGCGGACGGGCTCATCAAGTCGCTAAACCTCTGATGAAGAAGCAGCGGTGAAGCGC 261  
Db 195 TGGCGGAGGGCTGATCAAGTCGCCCCAAGCCCTTAATGAAGAAGCAGCGGTGAAGCGC 254  
Qy 262 ACCATCAAAACAACACCTCGCGCACCGCTACGAGTTCCTGGAGACGCTGGCAAGGGCA 321  
Db 255 ACCACCAAGCACAACTCGCGCACCGCTACGAGTTCCTGGAGACCTCGGCAAGGGCA 314  
Qy 322 CTTACGGGAAGGTGAAGAGGACGAGAGAGCTCGGGGCGCTCTGTTGGCCATCAAGTCCA 381  
Db 315 CTTACGGGAAGGTGAAGAGGCGCGGAGAGCTCGGGGCGCTCTGTTGGCCATCAAGTCCA 374  
Qy 382 TCAGGAAAGACAAATCAAGATGAGCAGGATCTGCTGCACATACGAGGAGATTGAGA 441  
Db 375 TCCGGAAGACAAATCAAGATGAGCAGGATCTGATGCACATACGAGGAGATTGAGA 434  
Qy 442 TCATGTTTCACTCAACCCACCATCATTCGCCATCCATGAAGTGTGTTGAGAATAGCA 501  
Db 435 TCATGTCATCACTCAACCCACCATCATTCGCCATCCATGAAGTGTGTTGAGAATAGCA 494  
Qy 502 GCAAGATTGTGATTGTGATGGAGTATGCCAGCGGAGCGATCTGTATGATTACATCAGTG 561  
Db 495 GCAAGATTGTGATTGTGATGGAGTATGCCAGCGGAGCGATCTGTATGATTACATCAGTG 554  
Qy 562 AGCGGCCACGGCTGAGTGAGCGGAGCGCCAGGCGATTTCTCCGACAGATCGTGTGCC 621  
Db 555 AGCGGCAGCAGCTCAGTGAGCGGAGCGATGAGGATTTCTCCGCGAGATCGTCTGCGCG 614  
Qy 622 TGCACTACTGCCACAGAGCGGATCGTTTCCACCGAGATCTCAAGCTGGAAAAACATCCTTC 681  
Db 615 TGCACTATTGCCATCAGACAGAGTGTGCCACCGAGATCTCAAGCTGGAGAACATCCTCT 674  
Qy 682 TAGATGCCATGGAACATCAAGATTGCTGACTTTGGCCCTCCACCTGTACCAAG 741  
Db 675 TGGATGCCAATGGGAATATCAAGATTGCTGACTTTGGCCCTCCACCTGTACCATCAAG 734  
Qy 742 GCAAGTTTCTCCAGAGCTTCTGTGGGAGCGCTCTCTACGCTCGCTCGATGATAGTCAACG 801  
Db 735 GCAAGTTTCTCCAGAGATTTCTGTGGAGCGCCCTCTATGCTCGCCAGAGATTGTCAATG 794  
Qy 802 GGAAGCCCTATGTGGGCGCCAGAGGTGGACAGCTGGTCTCTGGGCGTTCTCTGTACATCC 861  
Db 795 GGAAGCCCTACACAGGCGCCAGAGGTGGACAGCTGGTCCCTGGGTGTTCTCTCTACATCC 854  
Qy 862 TGGTGCATGSCACCATGCCCTTTGACGGGCGAGATCATAAACACTGGTGAAGCAATCA 921  
Db 955 TGGTGCATGSCACCATGCCCTTTGATGGGCGATGACCAATAGATCTCTAGTAGAAGATCA 914  
Qy 922 GTAACGGGGCTTACCGGTGAGCGCCCAAGCGCTCCGATGCTGTGGCTCATCGGTGGC 981  
Db 915 GCNAACGGGGCTTACCGGGAGCCACTTAACCTCTGTATGCTGTGGCTGATTCGGTGGC 974  
Qy 982 TGTTAATGTGTAACCCCGCTCGGGGCCACCTGGAGGATGTAGCCAGTCAATTTGGTGGG 1041

975	TGTTGATGGTGAA	1034	CCCGCCGGGCGCAC	1035	CCCTGGAGGATGTGGCCAGTCACTG	1036	TGGTGCGG
1042	TCAACTCGGGTTT	1037	ACCA	1038	CCGAGTCTGGGGAA	1039	CAGGAAGCCCTCGTGAGGGTGGGCACC
1035	TCAACTCGGGCT	1040	ACGCCACCGAGTGGGAGACGAGGAGTCCGCATGATGAGGGTGGGCACC	1094			
1102	CTAGTGTGTGACTT	1041	TGGCGGGCCCTCCATGGCGGACTGGTTACGTTCGCTCTCTCGGCGCCCC	1161			
1095	CTGGCACTGACTCT	1042	TGCGCGGCGCTCCCATGGCTGACTGGCTCCGCGCTTCTCTCCGCGCCCC	1154			
1162	TCCTGGAGAAAT	1043	TGGAACCAAGGTGTGCAGCTTCTTCAAGCAGCAGCTGCGCGGAGGTGCAA	1221			
1155	TCCTGGAGAAAT	1044	TGGGCCCAAGGTGTGCAGCTTCTTCAAGCAGCATGCACCTGGTGGGGAA	1214			
1222	GCACGTGACCTT	1045	TGGCTGGAGCGGCAATTTCTTTAAGAGTGTCCGAAAGAGAAATGACA	1281			
1215	GCACCACTCGGCT	1046	TGGAGCGCCAGCAATTCGCTCAAGAAATCCCGCAAGGAGAAATGACA	1274			
1282	TGGCTCAAAAT	1047	TGCAAGGTGACCCGCTGAGGATACCTTCTTCGCGCTGGCGAAGAGCA	1341			
1275	TGGCCAGTCTCT	1048	CCACAGTGACACGGCTGATGACATGCCCATTCGCCCTTGGCMAAGAGCA	1334			
1342	GCCTTAAGCTTC	1049	CGAAAGGCATTTCTCAAGAAAAAGTCTCTTACCTCTCAGGGGAGGTAC	1401			
1335	ACCTCAAGCTG	1050	CCAAAGGGCAATTTCTCAAGAAGAAAGTGTCAAGCTCTGCAAGAGGGGTAC	1394			
1402	AGGAGGACCTCT	1051	CAGGAACTCAGACCGGTGCTGATATCTCCAGGCGACGCTGTCCCTGCTG	1461			
1395	AGAGGACCTCT	1052	CGGAGCTCAGCCCAATTCCTGCGAGCCGAGGCGAGCTGCCCC	1448			
1462	TATCTCCTGCT	1053	CCCCAAGAAAGGCATCTCTTAAGAGTCTCGACAGCGTGAATCTGCTTACT	1521			
1449	---	1054	CGCTGCTCCCCAAGGGCAATTTCTCAAGAACCCCGACAGCGGAGTCTGGCTACT	1505			
1522	ACTCTCTCAGAG	1055	CCCGAGTCTGGGAACTCTTAGAGCGCCAGTGATGTGTTTGTA	1581			
1506	ACTCTCTCCGAG	1056	CCCAAGTAACTCTGGGAGCTCTTGGACGCGAGGCGAGTGTTTGTGA	1565			
1582	GTGGGACCCG	1057	TGGACAGAAGTCTCCAAGGCTTCAGGGCTCTCTCTCCACCGCAAGG	1641			
1566	GTGGGATCC	1058	AGGAGCAGAGCCCTCGCAGCTTCAGGGCTGCTCTCTCATCGCAAG	1625			
1642	GCATTTCAAAT	1059	CTAATGGCAAGTTTCTCCGCA	1600	CAGCCTTAGAAGGCATCTACCCCTAGCA	1701	
1626	GCATCTCAAAC	1060	TCAATGGCAAGTTTCTCCAGACAGCCTTGGAGCTCGCGGCCCCACCA	1685			
1702	CTTTGSGCTCC	1061	TGGACCAACTGSGCTCTCCCATCTGACGCGCGCCGACGCGCCCT	1761			
1686	CTTTGSGCTCC	1062	TGGATGAACTCGCCCCCACTCGCCCCCTTGGCCCCGGGCGACGCGACCT	1745			
1762	CAGGGCTGTG	1063	AGGACAGCATCTGTCTCCGAGTCTTTTGACCAATTTGCACTTGC	1821			
1746	CAGGGCTGTG	1064	AGGACAGCATCTGTCTCTCTGAGTCTTTTGACCACTGCACTTGC	1805			
1822	CTGAAGCTCT	1065	CCGAAACCCCACTGAGGGCTGTGTCTGTGGACAACTGAGGGGGC	1881			
1806	CTGAACGGCT	1066	CCAGAGCCCCCACTCGGGGCTGTGTGTCTGTGGACAACTCA	1865	CGGGGC		
1882	TTGAGCAGCCT	1067	CTCCTCAGAA-	1868	-----GTCTGAAGCGATGTGGCAGGAATCTCT	1929	
1866	TTGAGGAGCCCC	1068	CTCAGAGGGCCCTGGAAGCTGCCTGAGCGCTGCGCGAGGATCCTT	1925			
1930	TGGGGATAGCT	1069	GTCTTTCTGACAGACTGCGCAAGAGGTGACTGACGCTTACAGCAAG	1989			
1926	TGGGGACAGCT	1070	GTCTTCTGACAGACTGCGCAGAGGTGACAGCGACCTACCGACAGG	1985			
1990	CCCTAGGAAT	1071	CTGTCAAAGCTCAGCTGAGGAAGGAGATGGTGTCCCTAGTATGGGGTAG	2049			
1986	CATGTAGGGT	1072	CTGTCAAAGCTCACCTGATGTGAGTAGGCAATGTCCCCAG-CCCGGTCTAG	2044			
2050	GCTCTGAGAGGG	1073	TTTGAGAGGAACCTCGGGTCGGATTCCTCC	2092			
2045	GCTCTCAGATG	1074	CAGCTGGTGCACCCCGAGGGGAGATGCTTCC	2087			

148 CTAATTGATTCCTCCGCGCCCTTGTCTCACTCTCTGCTGCCATGAGTCGCTGTTTTCG 207  
142 TCACAGCCGCCAGCCAGCTCTCTCGCCTCCGCTCGGCTCGGAGAGCGCCCGCGCGC 201  
208 CGCGCGCTCCGCGCCCACTCCCTCGCGCGCAGAGCTA-----GCCCGCGCGC 255  
202 TGGCGGAGCGGCTCATCAAGTCGCTTAACTCTCTGATGAAGAAGCAGCGCGTGAAGCGGC 261  
256 TGGCGGAAGGCTGATCAAGTCGCTTAACTCTCTGATGAAGAAGCAGCGCGTGAAGCGGC 315  
262 ACCATCAAAACAACAACCTTCGCGCACCTCTACGATTCCTGAGACGCGTGGGCAAGGGCA 321  
316 ACCACCAAGCAACAACCTTCGCGCACCTCTACGATTCCTGAGACGCGTGGGCAAGGGCA 375  
322 CCTACGGGAAGGTGAAGAAGCAGCAGAGAGCTCGGGCGCTCTGTCGCCATCAAGTCCA 381  
376 CCTACGGGAAGGTGAAGAAGCAGCAGAGAGCTCGGGCGCTCTGTCGCCATCAAGTCCA 435  
382 TCAGGAAGCAAAATCAAGATGAGCAGGATCTGCTGCAATACGAGGAGGATTTGAGA 441  
436 TCCGGAAGGCAAAATCAAGATGAGCAGGATCTGATGACATACGAGGAGGATTTGAGA 495  
442 TCATGCTTCTCAACACCCCAATCAATGTCATGTCATGTCATGTCATGTCATGTCATGTC 501  
496 TCATGCTTCTCAACACCCCAATCAATGTCATGTCATGTCATGTCATGTCATGTCATGTC 555  
502 GCAAGATGATGATGTCATGAGTATGCCAGCGAGGATCTCTATGATACATCAGTG 561  
556 GCAAGATGATGATGTCATGAGTATGCCAGCGAGGATCTCTATGATACATCAGTG 615  
562 AGCGGCCACAGCTGAGTGAGCGGAGCGCAGGCAATTTCTCCGACAGATCGTCTGCGCC 621  
616 AGCGGCCACAGCTGAGTGAGCGGAGCGCAGGCAATTTCTCCGACAGATCGTCTGCGCG 675  
622 TGCACTACTGCCAAGAGCGGATCGTTCACGAGATCTCAAGCTGGAAGCAATCTTTC 681  
676 TGCACTACTGCCAAGAGCGGATCGTTCACGAGATCTCAAGCTGGAAGCAATCTTTC 735  
682 TAGATGCCAATGGAAGCAATCAAGATGTCGACTTTGGCTCTCAAGCTGGAAGCAATCAAG 741  
736 TGATGCCAATGGAAGCAATCAAGATGTCGACTTTGGCTCTCAAGCTGGAAGCAATCAAG 795  
742 GCAAGTTCCTCCAGACGCTTGTGGAGCGCTCTCTACGCTCGCCTCGAGATAGTCAACG 801  
796 GCAAGTTCCTCCAGACGCTTGTGGAGCGCTCTCTATGCTTCGCCAGAGATTTCAATG 855  
802 GGAAGCCTATGTTGGGCGCCAGAGTGAGCAGCTGCTCTGGGCGTTCCTCTGATATCC 861  
856 GGAAGCCTATGTTGGGCGCCAGAGTGAGCAGCTGCTCTGGGCGTTCCTCTGATATCC 915  
862 TGGTGCATGGCACCACCTGCTTTCAGCGGAGGATCATAAACACTGGTGAAGCAATCA 921  
916 TGGTGCATGGCACCACCTGCTTTCAGCGGAGGATCATAAACACTGGTGAAGCAATCA 975  
922 GTAACGGGCTTACCGTGAGCGCGCCCAAGCGCTCGGATCGCTGCTGCTGCTGCTGCTG 981  
976 GCAACGGGCTTACCGTGAGCGCGCCCAAGCGCTCGGATCGCTGCTGCTGCTGCTGCTG 1035  
982 TGTAAATGTTGAACCCCAACCGCTCGGCGCACACTGGAGGATGTAGCCAGTCAATTTGGTGG 1041  
1036 TGTAAATGTTGAACCCCAACCGCTCGGCGCACACTGGAGGATGTAGCCAGTCAATTTGGTGG 1095  
1042 TCAACTGGGTTACACCAAGGATCGGGGAACAGGAGCGCTCGTGAAGGTTGGGCAAC 1101  
1096 TCAACTGGGTTACACCAAGGATCGGGGAACAGGAGCGCTCGTGAAGGTTGGGCAAC 1155  
1102 CTAGTGGTGAATTTGGCGCGGCTTCCATGGCGGAGTGGTTACGTCGCTCTCTCGCGCCGCC 1161  
1156 CTGGCAGTGAATTTGGCGCGGCTTCCATGGCGGAGTGGTTACGTCGCTCTCTCGCGCCGCC 1215  
1162 TCCTGGAGATGGAGCCCAAGGTGTGCACTTTTCAAGCAGCAGCTGCGGAGGTTGGA 1221  
1216 TCCTGGAGATGGAGCCCAAGGTGTGCACTTTTCAAGCAGCAGCTGCGGAGGTTGGA 1275

QY 1222 GCATGTATCTGCTGGAGCGGCAACATTCTCTTAAGAAGTCCCGAAAGAGATGACA 1281  
Db 1276 GCACACCCCTGCGCTGGAGCGGCAATTCGCTCAAGAAGTCCCGAAGAGATGACA 1335  
QY 1282 TGCTCAAAATCTGCAAGGTGACCCCGCTGAGGATACCTCTTCTGCGCTTGGCAAGACA 1341  
Db 1336 TGGCCAGTCTCTCCACAGTGAACAGGCTGATGACATGCGCCATGCGCTTGGCAAGACA 1395  
QY 1342 GCCTTAAGCTTCCGAAGGCAATTCNAAGAAAGTCTCTACCTGCTCAGGGGAGTAC 1401  
Db 1396 ACCTCAAGCTGCCAAAGGCAATTCNAAGAAAGTGTGAGCTCTGCAAGAGGGGATAC 1455  
QY 1402 AGGAGACCTCTCAGGAACTCAGACCGGTGCTGATATCTCCAGGCGACCTGTCCCTGCTG 1461  
Db 1456 AGAGGACCTCTCGGAGCTCAGCCCAATCCCTGGAGCCACAGGCGAGCTTGCC- 1509  
QY 1462 TATCTCTGCTCCCAAGAAAGGCAATCTTAAGAAAGTCTCCACAGCGTGAATCTGGTTACT 1521  
Db 1510 ---CCCTGCTCCCAAGAGGCAATTCNAAGAAAGTGTGAGCTCTGCAAGAGGGGATAC 1566  
QY 1522 ACTCTCTCAGAGCCCGAGTCTGGGAACTCTTAGACGCGAGTGTGTTTGA 1581  
Db 1567 ACTCTCTCAGAGCCCGAGTGAATCTGGGAGCTCTTGGACGCGAGCGTGTGTTGA 1626  
QY 1582 GTGGGACCCGTTGGAGCAGAGTCTCCACAGGCTTCAGGCTCTCCTCCACGCAAGG 1641  
Db 1627 GTGGGATCCCAAGAGCAGAGCTCCCGAGCTTCAGGCTGCTCTCTCAGCAAG 1686  
QY 1642 GCATCTCAAACTCAATGGCAAGTTCCTCCGACAGCTTTAGAGGCACTACCCCTAGCA 1701  
Db 1687 GCATCTCAAACTCAATGGCAAGTTCCTCCGACAGCTTTGGAGCTCGCGGCCCAACCA 1746  
QY 1702 CTTTGGCTCCTGACCAACTGCGCTCTCCATCTGACGCGCCGCGCCGCGCCCT 1761  
Db 1747 CTTTGGCTCCTGAGTGAAGTGAATCTGCGCCACCTCGCGCCGCGCGCGAGCCCT 1806  
QY 1762 CAGGGCTGTGAGTGAAGCAGATCTGCTCCGAGTCTTTGACCAATTTGACTTGC 1821  
Db 1807 CAGGGCTGTGAGGAGCAGATCTGCTCTCTGATCTTTGACAGCTGAGCTTGC 1866  
QY 1822 CTGAAGCTTCTCCGAAACCCCACTGAGGGCTGTGTCTGTGGACAACTGAGGGGC 1881  
Db 1867 CTGAAGCTTCTCCGAAACCCCACTGAGGGCTGTGTCTGTGGACAACTGAGGGGC 1926  
QY 1882 TTGAGAGCTCTCTGAGAG-----GTCTGAAGCAGTGTGTCAGAGATCTCT 1929  
Db 1927 TTGAGAGCTCTCTGAGAGCTCTGAGAGCTGCTGAGGCGCTGCGCGAGATCTCT 1986  
QY 1930 TGGGGATAGCTCTCTTCTGACAGCTGCAAGAGTGTGCTGACGCTTACAGCAAG 1989  
Db 1987 TGGGGAGACGCTCTCTTCTGACAGCTGCAAGAGTGTGCTGACGCTTACAGCAAG 2046  
QY 1990 CCTTAGGAATCTGCTCAAGCTCAGCTGAGGAGGAGATGCTGCTCTAGTATGGGATG 2049  
Db 2047 CACTGAGGCTGCTCAAGCTCAGCTGAGTGGAGTAGGATTTGCCCGAG-CCCGGTCAG 2105  
QY 2050 GCTCTGAGAGGTTTTCAGAGGAACTCTGGGTGAGATTCCTCC 2092  
Db 2106 GCTCTCAGATGACGCTGTTGTCACCCCGAGGGAGATGCTCTTC 2148

## RESULT 12

ADM43851

ID : ADM43851 standard; cDNA; 2501 BP.

XX

XX ADM43851;

XX

DT 03-JUN-2004 (first entry)

XX

XX Novel human arginine-rich protein cDNA #215.

XX

KW ss; gene; human; arginine-rich protein; cancer; inflammation;

Genetic disorder.  
 XX Homo sapiens.  
 XX US2004053250-A1.  
 XX 18-MAR-2004.  
 XX 21-NOV-2002; 2002US-00302172.  
 XX 05-MAR-2001; 2001US-00799451.  
 XX 05-MAR-2002; 2002WO-US005095.  
 XX 20-AUG-2002; 2002US-00225251.  
 XX (TANG/) TANG Y T.  
 XX (XUEA/) XUE A.  
 XX (DRMA/) DRMANAC R T.  
 XX Tang YT, Xue A, Drmanac RT;  
 XX WPI; 2004-238579/22.  
 XX New isolated arginine-rich protein-like polynucleotides and polypeptides,  
 XX useful for diagnosing and/or treating conditions associated with aberrant  
 XX activity of the arginine-rich polypeptides, such as cancer and  
 XX inflammation.  
 XX Disclosure; SEQ ID NO 215; 51pp; English.  
 XX The invention relates to an isolated polynucleotide. The methods and  
 XX compositions of the present invention are useful for the diagnosis and/or  
 XX treatment of diseases or conditions associated with aberrant expression  
 XX or activity of the arginine-rich protein-like polypeptides, such as  
 XX cancer and inflammation. They can also be used in forensics, gene  
 XX mapping, identification of mutations responsible for genetic disorders,  
 XX and in assessing biodiversity. The present sequence represents a novel  
 XX human arginine-rich protein cDNA.  
 XX Sequence 2501 BP; 525 A; 788 C; 728 G; 460 T; 0 U; 0 Other;  
 XX  
 XX Query Match 47.3%; Score 1372.2; DB 12; Length 2501;  
 XX Best Local Similarity 81.6%; Pred. No. 2.4e-233;  
 XX Matches 1649; Conservative 0; Mismatches 338; Indels 34; Gaps 4;  
 XX  
 XX 82 CTACTGTGTCGCGACCCACCTCCACCTCGCGTCCCGCACCATGGAGTCGGTCCCTTAC 141  
 XX 65 CTATGTGATTCCTCGCGCCCTTCTGCTCCTCGCTCGGCATGGAGTCGGTGTTCG 124  
 XX 142 TCAGAGCGCGGAGCGAGGCTCCCTCGGCCTCCGCTCGGCCTCGGAGCGCGCGCGC 201  
 XX 125 CGCGGCGCTCGGCGCCCACTCCCTCGCGCGAGAGCTA-----GCCGCGCGC 172  
 XX 202 TGGCGGACGGGCTCATCAAGTCGGCTTAACCTCTGTATGAAGACGAGCGGTGAAGCGC 261  
 XX 173 TGGCGGAAGGGCTGATCAAGTCGCGCCCAAGCCCTTAATCAAGAGCAGCGCGTGAAGCGC 232  
 XX 262 ACCATCAACAACAACCTCGGCACCGCTACGAGTTCCTGGAGCGCTCGGCAAGGCA 321  
 XX 233 ACCACCAAGACCAACCTCGGCACCGCTACGAGTTCCTGGAGACCTCGGCAAGGCA 292  
 XX 322 CCTACGGGAAGGTGAAGAGGACGAGAGAGCTCGGGCGCTCTGGTGGCCATCAAGTCCA 381  
 XX 293 CCTACGGGAAGGTGAAGAGGCGCGGAGAGCTCGGGCGCTTGGTGGCCATCAAGTCAA 352  
 XX 382 TCAGGAAGACAAATCAAGATGAGCAGATCTGCTGACATACGAGGAGAGATTGAGA 441  
 XX 353 TCGGAAGAGGACAAATCAAGATGAGCAGATCTGATGCAATACGAGGAGAGATTGAGA 412  
 XX 442 TCATGTCTTCACTCAACACCCCACTATTCATTCGATCCATGAAGTCTTTGAGAATAGCA 501  
 XX 413 TCATGTCTTCACTCAACACCCCTCAGATCATTCATTCGATCCATGAAGTCTTTGAGACAGCA 472  
 XX 502 GCAAGATTGTGATGTCATGGAGTATGCCAGCGGAGCGGATCTGTATGATTACATCAGTG 561  
 XX  
 XX Db 473 GCAAGATCGTATCGTATGCGAGTATGCGAGCGGCGGACCTTTATGATACATCAGCG 532  
 XX Qy 562 AGCGGCCACGGCTGAGTGAAGCGGAGCGCCAGGCAATTTCTTCCGACAGATCGTGTGCCCC 621  
 XX Db 533 AGCGGACGAGCTCAGTGAAGCGGAGTATTTCTTCCGCGAGATCGTCTGCGCG 592  
 XX Qy 622 TGCACTACTGCGACGAGACGGGATGTTTACCGAGATCTCAAGCTGGGAAACATCCTTC 681  
 XX Db 593 TGCACTATTGCGCATCAGAACAGAGTTGTCCACCGAGATCTCAAGCTGGGAAACATCCT 652  
 XX Qy 682 TAGATGCCAATGAAACATCAAGATTTGCGCTTCTCAACCTCTCAACCTCTCAACCTCAAG 741  
 XX Db 653 TGATGCCAATGGAATATCAAGATTTGCGTCTCTCAACCTCTCAACCTCTCAACCTCAAG 712  
 XX Qy 742 GCAAGTTCTCCAGAGCTTCTGTGGAGCGCTCTCTACGCTCGCTCGCTGAGATAGTCAAG 801  
 XX Db 713 GCAAGTTCTCGCAGACATTTCTGTGGAGCGCTCTCTATGCTCGCGAGAGATTGTCAATG 772  
 XX Qy 802 GGAAGCCTATGTGGGCGGACGAGTGGACAGCTGGTCTCTGGGCGTTCTCTCTATCATCC 861  
 XX Db 773 GGAAGCCTTACACAGGCGGACGAGTGGACAGCTGGTCTCTGGGCGTTCTCTCTATCATCC 832  
 XX Qy 862 TGGTGCATGGCAGCAGTCCCTTTGACGGGAGGATCATAAACACCTGGTGAAGCAATCA 921  
 XX Db 833 TGGTGCATGGCAGCAGTCCCTTTGATGGGAGTACCATAGATCTCTAGTGAACAGATCA 892  
 XX Qy 922 GTAAACGGGCTTACCTGAGCGCGCCCAAGCGCTCCGATGCTGTGCGCTGATCCGCTGCG 981  
 XX Db 893 GCAACGGGCGCTACCGGAGGACCACTTAAACCTCTGATGCTGTGCGCTGATCCGCTGCG 952  
 XX Qy 982 TGTATATGTGAACCCACCGCTCGGGCCACATCGAGGATGTAGCCAGTCAATGCTGGG 1041  
 XX Db 953 TGTATATGTGAACCCACCGCTCGGGCCACATCGAGGATGTAGCCAGTCAATGCTGGG 1012  
 XX Qy 1042 TCAACTGGGGTTACACACCGGAGTCCGGGAAACAGGAAGCCCTGCTGAGGGTGGGCAAC 1101  
 XX Db 1013 TCAACTGGGGTTACGCGCACCCGAGTGGGAGAGCAGAGGCTCGCATAGGGTGGGCAAC 1072  
 XX Qy 1102 CTAGTGGTGAATTTGGCGCGGCTCCATGCGGAGTCTGTTTACGCTGCTCTCGCGGCCCC 1161  
 XX Db 1073 CTGGCAGTGAATCTGCGCGCGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1132  
 XX Qy 1162 TCGTGCAGATGAGCAGCAAGGTGTGAGTCTTCAAGCAGACAGCTGCGGAGGAGTGGAA 1221  
 XX Db 1133 TCGTGCAGATGAGCAGCAAGGTGTGAGTCTTCAAGCAGACAGCTGCGGAGGAGTGGAA 1192  
 XX Qy 1222 GCACTGTACTTGGGCTGGAGCGGCAACATTTCTTAAAGAGTCCCGAAAGGAGAAATGACA 1281  
 XX Db 1193 GCACCAACCTTGGCTGGAGCGGCAACATTTCTTAAAGAGTCCCGAAAGGAGAAATGACA 1252  
 XX Qy 1282 TGGCTCAAAATCTGAAGGTGACCGCGCTGAGGATACCTTCTGCTGCTGCTGCTGCTGCTGCT 1341  
 XX Db 1253 TGGCGCAGTCTCTCCACAGTGAACAGGCTGATGACATGCGCTGCTGCTGCTGCTGCTGCTGCT 1312  
 XX Qy 1342 GCCTTAAGCTTCCGAAGGCAATTTCTCAAGAAAGTCTTCAAGCAGACAGCTGCTGCTGCTGCTGCT 1401  
 XX Db 1313 ACCTCAAGCTGCCAAGGCAATTTCTCAAGAAAGTGTGAGCTTCTGAGAAAGGAGTAC 1372  
 XX Qy 1402 AGGAGGACCTTCAAGGAACTCAGACCGGCTGCTGATTAAGTCTGAGGAGCGCTGCTGCTGCTGCT 1461  
 XX Db 1373 AGGAGGACCTTCCGAGGCTCAGCCCAATCCCTGCGAGCGCGGAGGCTGCTGCTGCTGCTGCT 1426  
 XX Qy 1462 TATCCCTGCTCCCAAGGAAAGGCACTTCAAGAGTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCT 1521  
 XX Db 1427 ---CCCTGCTGCTCCCAAGGAGGCAATTTCTCAAGAAAGGCTGCTGCTGCTGCTGCTGCTGCT 1483  
 XX Qy 1522 ACTCTCTCCAGAGCGGAGTCTGGGAACTCTTTAGACGCGAGTGTGTTGTGA 1581  
 XX Db 1484 ACTCTCTCCAGAGCGGAGTGTGAGTCTGGGAGGCTCTTGAAGCGGAGGAGTGTGTTGTGA 1543  
 XX Qy 1582 GTGGGAGCGCGGTGGAGAGAGAGTCTCCACAGGCTTCAAGGCTTCAAGGCTTCTCTCCACCGCAAG 1641



Db 1544 GTGGGATCCCAAGGAGCAGAAGCCTCCGCAAGCTTCAGGGCTGCTCTCCATCGCAAG 1603  
Qy 1642 GCATTCTCAAACTCAATGGCAAGTTCTCCGCGCAGACCTTAGAGGCACTACCCCTAGCA 1701  
Db 1604 GCATCTCAAACTCAATGGCAAGTTCTCCGCGCAGACCTTAGAGGCTCGGGCCCCACCA 1663  
Qy 1702 CCTTTGGCTCCCTGGACCAACTGSCCTCTCTCCCATCTCGACGCCCGCCAGCGCCCT 1761  
Db 1664 CCTTCGGCTCCTGGATGAATCGSCCCACCTCGCCCTCGGGCCGCGCAGCCCT 1723  
Qy 1762 CAGGGCTGTGAGTGAGGACAGCATCTGTCTCCGAGTCTTTGACCAATGAGCTTGC 1821  
Db 1724 CAGGGCTGTGAGTGAGGACAGCATCTGTCTCCGAGTCTTTGACCAATGAGCTTGC 1783  
Qy 1822 CTGAAGCTCTCCGGAACCCCACTGAGGGGCTGTGTCTGTGACAACTGAGGGGC 1881  
Db 1784 CTGACGGCTCCGAGACCCCACTGCGGGCTGTGTCTGTGACAACTGAGGGGC 1843  
Qy 1882 TTGAGCAGCTCCCTCAGAAG-----GTCTGAAGCGATGTGGCAGGAATCCT 1929  
Db 1844 TTGAGGAGCCCTCAGAGGGCCCTGAAGCTGCTGAGCGCTGCGCGAGGATCCTT 1903  
Qy 1930 TGGGGATAGTGTCTTTCTGTGACAGACTGCCAAGAGGTGACTGCGACCTACAGCAAG 1989  
Db 1904 TGGGGGACAGTGTCTTTCTGTGACAGACTGCCAAGAGGTGACAGCGACCTACCGACAG 1963  
Qy 1990 CCCTAGGAATCTGCTCAAGCTCAGCTGAGAGGGAGATGGTCCCTAGTATGGGTAG 2049  
Db 1964 CACTGAGGGTGTCTCAAGCTCAGCTGAGTGGAGTAGGCAATGCCCCAG-CCGGGTAG 2022  
Qy 2050 GCTCTGAGAGGGTTTGCAGAGGAACCTCGGTGCGATTCCT 2090  
Db 2023 GCTCTCAGATGCAGTGTGTTGACCCCGAGGGGAGATGTCCT 2063

RESULT 13  
ABZ11333  
ID ABZ11333 standard; cDNA; 2043 BP.  
XX  
AC ABZ11333;  
XX  
DT 20-JAN-2003 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 215.  
XX  
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;  
KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;  
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;  
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
KW antiarthritic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200270539-A2.  
XX  
PD 12-SEP-2002.  
XX  
PF 05-MAR-2002; 2002WO-US005095.  
XX  
PR 05-MAR-2001; 2001US-00799451.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
PI Wehrman T, Wang J, Wang D, Drmanac RT;  
XX  
DR WPI; 2002-759812/82.  
DR P-PSDB; ABP69116.  
XX

PT New polynucleotides comprising sequences assembled from expressed  
PT sequence tags (ESTs), useful for treating cell-proliferative,  
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
PT or coagulation disorders.  
PS Claim 1; SEQ ID NO 215; 1012pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated polynucleotide (I) comprising a  
CC nucleotide sequence selected from any of 948 sequences (ABZ11119-  
CC ABZ12066) or their mature protein coding portion, active domain coding  
CC protein or complementary sequences. The polynucleotides are useful for  
CC identifying expressed genes or for physical mapping of human genome. The  
CC encoded polypeptides (ABP69849) are useful as molecular weight  
CC markers, as a food supplement, for generating antibodies, in medical  
CC imaging, screening and diagnostic assays and for treating cell-  
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,  
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver  
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
CC arthritis, etc. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 2043 BP; 418 A; 659 C; 592 G; 374 T; 0 U; 0 Other;  
Query Match 47.3%; Score 1371.8; DB 6; Length 2043;  
Best Local Similarity 82.3%; Pred. No. 2.8e-293;  
Matches 1625; Conservative 0; Mismatches 317; Indels 33; Gaps 3;  
Qy 82 CTACTGCTGCCCGACCCACTCCACCTCGCGGTCCCGCACCATCGAGTCGGTGGCTTAC 141  
Db 65 CTATTGATTCCCTGCGCCCTTGCTCACTCTGCTCGCATGAGTCGCTGGTTTCG 124  
Qy 142 TCCAGCGCCGACGAGGCTCCTCGGCTTCGCCCTCGGCTCGGAGAGCGCCGCGCGC 201  
Db 125 CGCGGGCTCCGGCCCACTCCCTCGCGCGCAGAGCTA-----GCCCGCGCGC 172  
Qy 202 TGGCGGACGGCTCATAGTCGCTTAACTCTGATGAGAGACAGCGCGTGAAGCGGC 261  
Db 173 TGGCGGAGGGCTGATCAAGTCGCCCAAGCCCTTAATGAAGAAGAGCGCGGTGAAGCGGC 232  
Qy 262 ACCATCACAAACAAACCTCGCGCACCGCTACGAGTTCCTGGAGAGCGCTGGGAAGGGCA 321  
Db 233 ACCACACACACACACCTCGCGCACCGCTACGAGTTCCTGGAGACCTCGGCAAGGGCA 292  
Qy 322 CTTACGGGAAGGTGAAGAAGGCAAGAGAGCTCGGGGCGCTCTGGTGCCATCAAGTCCA 381  
Db 293 CTTACGGGAAGGTGAAGAAGGCGCGGAGAGCTCGGGGCGCTCTGGTGCCATCAAGTCCA 352  
Qy 382 TCAGGAAGACAAATCAAGATGAGCAGGATCTGCTGCACATACGAGGGGAGATTGAGA 441  
Db 353 TCCGGAAGGACAAATCAAGATGAGCAGGATCTGATGCACATACGAGGGGAGATTGAGA 412  
Qy 442 TCATGTCTTCACTCAACCCACCTCCATTCCTCCATGAGTGTGTTGAGAATAGCA 501  
Db 413 TCATGTCTTCACTCAACCCACCTCCATTCCTCCATGAGTGTGTTGAGAATAGCA 472  
Qy 502 GCAAGATTGTGATTGTGATGAGTATGCGAGCGGAGCGGATCTGTATGATATCATCATG 561  
Db 473 GCAAGATTGTGATTGTGATGAGTATGCGAGCGGAGCGGATCTGTATGATATCATCATG 532  
Qy 562 AGCGGCCACGGCTGAGTGAGCGGAGCGGAGGATTTCTTCCGACAGATCGTCTGCGCC 621  
Db 533 AGCGGACGAGCTCAGTGAGCGGAGCGGAGGATTTCTTCCGCGAGATCGTCTGCGCG 592  
Qy 622 TGCACTACTGCCACCAAGACGGGATCGTTCAACCGAGATCTCAAGCTGGAAACATCCTTC 681  
Db 593 TGCACTACTGCCCATCAGACAGAGTTGTCCACCGAGATCTCAAGCTGGAGAACATCCTCT 652  
Qy 682 TAGATGCCAATGGAACATCAAGATTGCTGACTTTGCGCTCTCCAACTCTACCAAG 741  
Db 653 TGGATGCCAATGGAATATCAAGATTGCTGACTTTGCGCTCTCTCCAACTCTACCAATCAAG 712



QY 742 GCAAGTTCCTCCAGAGCTTCTGTGGAGCCCTCTCTACGCTCGCTGAGATAGTCAACG 801  
Db 713 GCAAGTTCCTCCAGAGCTTCTGTGGAGCCCTCTCTATGCTTCGCGAGAGATTGTCAATG 772  
QY 802 GGAAGCCCTATGTGGGCCAGAGGTGGACAGCTGTCTCTGGGGTCTCTCTGTACATCC 861  
Db 773 GGAAGCCCTACACAGGCCACAGGTGGACAGCTGTCTCTGGGTCTCTCTCTACATCC 832  
QY 862 TGGTGCATGGACCATGCTCTTTGACGGGAGGATCATATAACACTGGTGAAGCAAAATCA 921  
Db 833 TGGTGCATGGACCATGCTCTTTGATGGGATGATGACCATTAAGATCTTAGTGAACAGATCA 892  
QY 922 GTAACGGGCTTACCTGTAGCGCCCAAGCCGTCGATGCTGTGGCTGATCGGTGGC 981  
Db 893 GCAACGGGCTTACCGGGAGCCACTTAACCCCTCTGTATGCTGTGGCTGATCGGTGGC 952  
QY 982 TGTAAATGGTGAACCCCAACCCGTCGGGCCACACTGGAGGATGTAGCCAGTCAATTTGGTGG 1041  
Db 953 TGTGATGGTGAACCCCAACCCGTCGGGCCACCCCTGGAGGATGTGCCAGTCACTTGGTGG 1012  
QY 1042 TCAACTGGGTTACACACCGGATGCGGGGAACAGGAAGCCCTCGTGAGGTTGGGCACC 1101  
Db 1013 TCAACTGGGTTACACACCGGATGCGGGGAACAGGAAGCCCTCGCATGAGGTTGGGCACC 1072  
QY 1102 CTAGTGTGACTTTGGCGGGCCCTCCATGGCGGACTGTTACGTCGCTCTCGCGGCCCC 1161  
Db 1073 CTGCACTGACTCTGCGCGGCCCTCCATGGCTGACTGCTCGCGGTCTCTCGCGGCCCC 1132  
QY 1162 TCCTGGAGATGGAGCCAAAGGTGTGCAGCTTCTTCAAGCAGCAGCTGCGCGGAGGTGAA 1221  
Db 1133 TCCTGGAGATGGAGCCAAAGGTGTGCAGCTTCTTCAAGCAGCAGCTGCGCGGAGGTGAA 1192  
QY 1222 GCATCTACCTGGCTGGAGGGCAACATTTCTTAAGAAGTCCCGAAGAGGAATGACA 1281  
Db 1193 GCACACCTCTGGCTGGAGGGCCAGCATTCGCTCAAGAAGTCCCGAAGAGGAATGACA 1252  
QY 1282 TGGCTCAAAATCTCAAGGTGACCCGGCTGAGGATACCTCTTTCGCGCTGGCAAGACA 1341  
Db 1253 TGGCCAGTCTCTCCAGTGAACAGCGTGTATGACCTGCCCATCGCCCTGGCAAGACA 1312  
QY 1342 GCCTTAAGTCTCGAAAGGCAATCTCAAGAAAAAGTCTCTACCTCTCGAGGGAGGTAC 1401  
Db 1313 ACCTCAAGTCTCGAAAGGCAATCTCAAGAAAGAGGTCTCAGCCTCTCGAAGAGGGTAC 1372  
QY 1402 AGGAGGACCTTCAGGAATCAGACGGTGTGCTGATCTCAGGGCAGCTGTCTCTGCTG 1461  
Db 1373 AGGAGGACCTTCAGGAATCAGACGGTGTGCTGATCTCAGGGCAGCTGTCTCTGCTG 1426  
QY 1462 TATCCCTGCTCCCAAGGAAGGCAATCTTAAGAAGTCTCGACAGCGTGAATCTGTTACT 1521  
Db 1427 ---CCCTGCTCCCAAGGAAGGCAATCTTAAGAAGTCTCGACAGCGTGAATCTGTTACT 1483  
QY 1522 ACTCTCTCCAGAGCCCGAGGAGTCTGGGAACTCTTTAGAGCGCAGTGAATGTTGTGA 1581  
Db 1484 ACTCTCTCCAGAGCCCGAGTGAATCTGGGAGCTTTGGAGCGCAGGAGCTGTTGTGA 1543  
QY 1582 GTGGGACCCGCTGGAGCAGAGTCTCAAGGCTTCAGGGCTCTCTCTCAGCGAAG 1641  
Db 1544 GTGGGATCCCAAGGAGCAGAGCTCTCGAAGCTTCAGGGCTCTCTCTCAGCGAAG 1603  
QY 1642 GCATTTCTCAAACTCAATGGCAAGTCTCCGCGCAGAGCTTAGAAGGCACTACCCCTAGCA 1701  
Db 1604 GCATTTCTCAAACTCAATGGCAAGTCTCCGCGCAGAGCTTAGAAGGCACTACCCCTAGCA 1663  
QY 1702 CTTTGGCTCTCTGGACCAATGCTGCTCTCTCCATCTGAGCCCGCCAGCGGCCCT 1761  
Db 1664 CTTTGGCTCTCTGGATGAATCTGCCCCCACTCTGCCCTCTGGCCCGGCGCAGCGACCT 1723  
QY 1762 CAGGGGCTGTGAGTGAAGCAGCATCTGCTCCGAGTCTTTGACCAATTGACTTGC 1821  
Db 1724 CAGGGGCTGTGAGGAGCAGCATCTGCTCTCTGCTCTGAGTCTTTGACCACTGACTTGC 1783

QY 1822 CTGAACGCTCTCCGAAACCCCACTGAGGGGCTGTGTCTGTGGACAACTGAGGGGGC 1881  
Db 1784 CTGAACGCTCTCCGAAACCCCACTGAGGGGCTGTGTCTGTGGACAACTTACGGGGC 1843  
QY 1882 TTGAGCAGCTCTCCCTCAGAAG-----GTCTGAAGCGATGTGTGCGAGGAATCCT 1929  
Db 1844 TTGAGGAGCCCTCTCAGAGGGCCCTGGAAGCTGCTTGAAGCGTGTGCGGAGGATCCTT 1903  
QY 1930 TGGGGGATAGCTGCTTTTCTTCTGACAGACTGCCAAGAGGTGACTGACGCTTACAGACAAG 1989  
Db 1904 TGGGGGACAGCTGCTTTTCTTCTGACAGACTGCCAAGAGGTGACAGCGACCTTACGACAGG 1963  
QY 1990 CCTAGCAATCTGCTCAAGCTCAGCTGAGAGGAGAGATGCTGCCCTAGTATGG 2044  
Db 1964 CACTGAGGGTCTGCTCAAGCTCAGCTGAGTGGAGTAGGCAATGCCCCAGCCCGG 2018

RESULT 14  
AAD26459  
ID AAD26459 standard; cDNA; 3360 BP.  
XX AAD26459;  
XX AC  
XX DT 26-MAR-2002 (first entry)  
XX XX  
XX DE Human kinase PKIN-12 cDNA.  
XX KW Human; kinase; PKIN-12; cancer; leukaemia; adenocarcinoma; osteoporosis;  
KW immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;  
KW Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;  
KW allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;  
KW autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis;  
KW Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;  
KW rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;  
KW hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;  
KW cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;  
KW congestive heart failure; ischaemic heart disease; lung tumour; gout;  
KW fatty liver; Niemann-Pick's disease; gene therapy; ss.  
XX OS Homo sapiens.  
XX XX  
XX FH Key Location/Qualifiers  
XX CDS 181..1965  
XX FT /\*tag= a  
XX FT /product= "Human PKIN-12 protein"  
XX XX  
XX PN WO200196547-A2.  
XX XX  
XX PD 20-DEC-2001.  
XX XX  
XX PF 14-JUN-2001; 2001WO-US019444.  
XX XX  
XX PR 15-JUN-2000; 2000US-0212073P.  
XX PR 23-JUN-2000; 2000US-0213467P.  
XX PR 30-JUN-2000; 2000US-0215651P.  
XX PR 07-JUL-2000; 2000US-0216405P.  
XX PR 13-JUL-2000; 2000US-0218372P.  
XX PR 25-AUG-2000; 2000US-0228056P.  
XX XX  
XX PA (INCY-) INCYTE GENOMICS INC.  
XX XX  
XX PI Yue H, Lal P, Bandnan O, Borowsky ML, Au-Young J, Lu Y;  
PI Gandhi AR, Tribouley CM, Walia NK, Yao MG, Lu DM, Greenwald SR;  
PI Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang Y;  
PI Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;  
PI Lo TP, Khan F, Recipon SA, Azimzai Y, Policky JL, Ding L;  
PI Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;  
XX WPI: 2002-090207/12.  
XX P-PSDB; AAE16266.  
XX XX  
XX PT New polypeptides, useful for diagnosing, treating or preventing disorders  
of growth and development, cardiovascular and lipid, and diseases such as







*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 11:04:39 ; Search time 8911.77 Seconds  
(without alignments)  
11866.113 Million cell updates/sec

Title: US-09-980-464-4  
Perfect score: 2902  
Sequence: 1 cactagtgcacaaagaat.....aaaaaaaaaaaaaaaaaaaaa 2902

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hlc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2865.8	98.8	2899	3 AK004737	AK004737 Mus muscu
2	2865.8	98.8	3098	3 AK033672	AK033672 Mus muscu
3	2804.4	96.6	2869	3 AK034082	AK034082 Mus muscu
4	1776	61.2	1896	9 AY399041	AY399041 Mus muscu
5	1264	43.6	1887	9 AY399039	AY399039 Homo sapi
6	951.4	32.8	1887	9 AY399040	AY399040 Pan trogl
7	721	24.8	746	6 B1653092	B1653092 603300616
8	720	24.8	795	6 CA319312	CA319312 UI-M-FW0-
9	714	24.6	727	7 CF729223	CF729223 UI-M-FW0-
10	700.2	24.1	932	4 BG915967	BG915967 602815231
11	688	23.7	705	7 CF726196	CF726196 UI-M-GZ0-
12	683.6	23.6	726	6 CB248251	CB248251 UI-M-FD0-
13	672	23.2	672	7 CO424322	CO424322 UI-M-HU0-
14	667.4	23.0	716	7 CO041415	CO041415 UI-M-FW0-
15	663	22.8	986	7 CF584809	CF584809 AGENCOURT
16	660.8	22.8	816	4 B1143828	B1143828 602907107
17	655.4	22.6	704	7 CO041178	CO041178 UI-M-FW0-
18	654	22.5	893	2 BF785290	BF785290 602108490
19	653.2	22.5	713	2 BB625283	BB625283 602108490
20	652.4	22.5	655	7 CK781308	CK781308 UI-M-G10-
21	651.6	22.5	805	4 EG174288	EG174288 602334589
22	646	22.3	646	4 BG694881	BG694881 NISC_IV09
23	645.8	22.3	948	2 BF789245	BF789245 602105144
24	639.4	22.0	722	4 B1143493	B1143493 602907673

25	639	22.0	669	2	BB078728	BB078728
26	627.6	21.6	782	7	CN261003	CN261003 170004247
27	626	21.6	939	2	BE912458	BE912458 60166074
28	624.6	21.5	1069	4	BM927376	BM927376 AGENCOURT
29	615.6	21.2	1022	5	BQ062868	BQ062868 AGENCOURT
30	608	21.0	608	7	CK625537	CK625537 mj10904.Y
31	600.8	20.7	616	4	B1647932	B1647932 603276014
32	599.4	20.7	660	7	CO043728	CO043728 UI-M-FD0-
33	596.2	20.5	1005	5	BQ060729	BQ060729 AGENCOURT
34	596.2	20.5	1029	5	BQ059508	BQ059508 AGENCOURT
35	596	20.5	674	5	BQ445704	BQ445704 UI-M-ERO-
36	593	20.4	600	4	BG806175	BG806175 2021-17 M
37	593	20.4	673	5	BQ746210	BQ746210 UI-M-ERO-
38	584.2	20.1	1012	5	BQ064231	BQ064231 AGENCOURT
39	582.8	20.1	818	7	CK483904	CK483904 AGENCOURT
40	562	19.4	1042	4	BM909401	BM909401 AGENCOURT
41	558.8	19.3	667	4	B1648260	B1648260 603278202
42	558.6	19.2	865	4	BG172224	BG172224 602333366
43	556	19.2	556	6	BY704937	BY704937 BY704937
44	556	19.2	556	6	CA535696	CA535696 C0222H05-
45	554.8	19.1	558	8	BH116351	BH116351 RPCI-24-2

## ALIGNMENTS

RESULT 1  
AK004737  
LOCUS

DEFINITION

AK004737 2899 bp mRNA linear HTC 03-APR-2004  
Mus musculus adult male lung cDNA, RIKEN full-length enriched  
library, clone:120003B22 product:weakly similar to PROBABLE  
SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.1-) [Homo  
sapiens], full insert sequence.

ACCESSION

AK004737.2 GI:26334437

VERSION

HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Mammalia; Eutheria; Chordata; Vertebrata; Euteleostomi;

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

**AUTHORS** The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
**TITLE** Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
**JOURNAL** Nature 420, 563-573 (2002)  
**REFERENCE** 6 (bases 1 to 2899)  
**AUTHORS** Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanganaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

**TITLE** Direct Submission  
**JOURNAL** Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
**COMMENT** On Dec 10, 2002 this sequence version replaced gi:12836134. Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGCGCGCACTCGAGTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGAGAGCGCGCACTCGAGTTTCTTTTCTTTTNN 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOUR.

**FEATURES**  
 source  
 Location/Qualifiers  
 1..2899  
   /organism="Mus musculus"  
   /mol\_type="cDNA"  
   /strain="C57BL/6J"  
   /db\_xref="FANTOM DB:1200013B22"  
   /db\_xref="taxon:10090"  
   /clone="1200013B22"  
   /sex="male"  
   /tissue\_type="lung"  
   /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
   /dev\_stage="adult"  
   94..1989  
   /protein\_id="BAB23518.1"  
   /db\_xref="GI:12836135"  
   /translation="MESVALLOPSPQAPSAALASARPLADGLIKSPKPLMKKQAV  
 KRHHKHNLRHVEFLTGKTYGVKKARESGRLVAIKSRKDKIKDEQDLHLIR  
 REIEMSSLNPHIILIHVEFNSKIVIMEYASRGDLYDIISERPLSERDARHFF  
 RQIVSAHYCHQNGIVHRDLKLEINLLDANGIKIADFGSLNLYHKGFLQTFCSPL  
 YASPEIVNGKYPVGVPSWLSGLLYILVHGTMFPDQDQDKTLVQIISNGAYREP  
 PSDGLIRLWLLMTRATLDVAHWVWVNGYTTGVGEQALREGHPGSGDFQRA  
 SMADLRRSRPLLEAKVCFKOHVPVGGSTVPLGRHSLKRSKENDMAONLQ  
 GPAPEDTSRPGKSLKPKGILKKSSSTSGEVEDPOELRPVPTPCQVPANSL  
 PKGLILKSRQESGYISPERSEGLDADSVFVSGDFVQKSPQASGLLHRRGI  
 LKNGKFSKTALEGTTPSTFGSLDQASHPARPSPGAVSEDSILSESFQDL  
 PERLPETPLRGCVSDNLRLEQPPSEGLKRWQESGLSGSCFSLDQCQVTAAYRQAL  
 GICSKLS"  
 CDS

polyA\_signal 2876..2881  
 /note="putative"  
 polyA\_site 2899  
 /note="putative"  
 ORIGIN  
 Query Match 98.8%; Score 2865.8; DB 3; Length 2899;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2867; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 31 GCCTGCTCGGGTCCGGCTGTGACTCTGAGCCCGCGCTCAGCGCGGCTGCTACTCTG 90  
 Db 2 GCCTGCTCGGGTCCGGCTGTGACTCTGAGCCCGCGCTCAGCGCGGCTGCTACTCTG 61  
 QY 91 CCCGACCCACTCCACTCCGCGTCCCGCAACCATGAGTGGTGGCTTACTCAGCGCC 150  
 Db 62 CCCGACCCACTCCACTCCGCGTCCCGCAACCATGAGTGGTGGCTTACTCAGCGCC 121  
 QY 151 CGAGCCAGGCTCCCTCGGCTCCGCTCGGCTCGGAGAGCGCCCGGCTGGCGAGC 210  
 Db 122 CGAGCCAGGCTCCCTCGGCTCCGCTCGGCTCGGAGAGCGCCCGGCTGGCGAGC 181  
 QY 211 GGCTCATCAAGTCGCTTAAACCTCTCATGAAGAAGCAGCGGCTGAAGCGGCACCATCACA 270  
 Db 182 GGCTCATCAAGTCGCTTAAACCTCTCATGAAGAAGCAGCGGCTGAAGCGGCACCATCACA 241  
 QY 271 AACCAACCTCGGCAACCGCTACGAGTTCCTGGAGACGCTGGGCAAGGCGACCTACGGGA 330  
 Db 242 AACCAACCTCGGCAACCGCTACGAGTTCCTGGAGACGCTGGGCAAGGCGACCTACGGGA 301  
 QY 331 AGGTGAAGAGGACACAGAGAGCTCGGGGCTCTGCTGGCCATCAAGTCCATCAGGAAG 390  
 Db 302 AGGTGAAGAGGACACAGAGAGCTCGGGGCTCTGCTGGCCATCAAGTCCATCAGGAAG 361  
 QY 391 ACAAAATCAAAATGAGCAGGAGTCTCTGCACATAGCGAGGAGATTGAGATCATGTCTT 450  
 Db 362 ACAAAATCAAAATGAGCAGGAGTCTCTGCACATAGCGAGGAGATTGAGATCATGTCTT 421  
 QY 451 CACTAACCAACCCCAATCATTTGCCATCCATGAAGTGTGTGAGATAGCAGCAAGATTG 510  
 Db 422 CACTAACCAACCCCAATCATTTGCCATCCATGAAGTGTGTGAGATAGCAGCAAGATTG 481  
 QY 511 TGATTGTCATGAGTAGTGCAGCGAGGCGATCTGTATGATTACATCAGTGAGCGGCCAC 570  
 Db 482 TGATTGTCATGAGTAGTGCAGCGAGGCGATCTGTATGATTACATCAGTGAGCGGCCAC 541  
 QY 571 GGCTGAGTGAGCGGAGCGCGAGCAATTTCTTCGACAGATCGTGTCTGCCCTGCATCT 630  
 Db 542 GGCTGAGTGAGCGGAGCGCGAGCAATTTCTTCGACAGATCGTGTCTGCCCTGCATCT 601  
 QY 631 GCCACAGAACGGGATCGTTCAACGAGATCTCAAGCTGGAACACATCTTCTAGATGCCA 690  
 Db 602 GCCACAGAACGGGATCGTTCAACGAGATCTCAAGCTGGAACACATCTTCTAGATGCCA 661  
 QY 691 ATGGAACATCAAGATTGTCTGACTTTGGCTCTTCCAACTGTACCAACCAAGGCAAGTTCC 750  
 Db 662 ATGGAACATCAAGATTGTCTGACTTTGGCTCTTCCAACTGTACCAACCAAGGCAAGTTCC 721  
 QY 751 TCAGACGTTCTGTGGGAGCCCTCTCTACGCTCGCTGAGATAGTCAACGGAAGCCCT 810  
 Db 722 TCAGACGTTCTGTGGGAGCCCTCTCTACGCTCGCTGAGATAGTCAACGGAAGCCCT 781  
 QY 811 ATGTGGGCCAGAGGTGGACAGCTGGTCTCTGGGCGTTCTCTGTATCATCTTGGTGCATG 870  
 Db 782 ATGTGGGCCAGAGGTGGACAGCTGGTCTCTGGGCGTTCTCTGTATCATCTTGGTGCATG 841  
 QY 871 GCACCATGCTTTGACGGGAGGATCATAAACACTGGTGAAGCAAAATCAGTAAACGGG 930  
 Db 842 GCACCATGCTTTGACGGGAGGATCATAAACACTGGTGAAGCAAAATCAGTAAACGGG 901  
 QY 931 CTTACCGTGAGCGGCCCAAGCCGTCGAGTCTGTGGCTGTATCCGGTGGCTGTTTAATGG 990  
 Db 902 CTTACCGTGAGCGGCCCAAGCCGTCGAGTCTGTGGCTGTATCCGGTGGCTGTTTAATGG 961



QY	991	TGAACCCACCCGTCGGGCCACACTGAGGAGTGTAGCCAGTCAATTGGTGGGTCAACTGGG	1050
Db	962	TGAACCCACCCGTCGGGCCACACTGAGGAGTGTAGCCAGTCAATTGGTGGGTCAACTGGG	1021
QY	1051	GTTTACACACCGAGTCGGGGAAACAGGAAGCCCTGCTGAGGGTGGGCAACCTAGTGGTG	1110
Db	1022	GTTTACACACCGAGTCGGGGAAACAGGAAGCCCTGCTGAGGGTGGGCAACCTAGTGGTG	1081
QY	1111	ACTTTGSCCGGGCTCCATGCGGACCTGTTAGTCTGCTCCTCGCGCCCTCTCTGGAGA	1170
Db	1082	ACTTTGSCCGGGCTCCATGCGGACCTGTTAGTCTGCTCCTCGCGCCCTCTCTGGAGA	1141
QY	1171	ATGGAGCAAGGTGTGAGCTTCTTCAAGCAGCAGCTGCGCGGAGGTGGAGCACTGTAC	1230
Db	1142	ATGGAGCAAGGTGTGAGCTTCTTCAAGCAGCAGCTGCGCGGAGGTGGAGCACTGTAC	1201
QY	1231	CTGGGCTGAGCGGGCAACATCTCTTAAAGAGTCCCGAAAGAGAAATGACATGGCTCAAA	1290
Db	1202	CTGGGCTGAGCGGGCAACATCTCTTAAAGAGTCCCGAAAGAGAAATGACATGGCTCAAA	1261
QY	1291	ATCTGAAGGTGACCGGCTGAGGATACCTTCTCGCCCTGCGCAAGAGCAGCTTAAGC	1350
Db	1262	ATCTGAAGGTGACCGGCTGAGGATACCTTCTCGCCCTGCGCAAGAGCAGCTTAAGC	1321
QY	1351	TTCCGAAAGGCATTCTCAAGAAAAGTCTCTACCTCGTCAGGGGAGGTACAGGAGACC	1410
Db	1322	TTCCGAAAGGCATTCTCAAGAAAAGTCTCTACCTCGTCAGGGGAGGTACAGGAGACC	1381
QY	1411	CTCAGGAACCTCAGACCGGTCCTGATATCTCCAGGGCAGCCTGCTGCTGTATCCCTGC	1470
Db	1382	CTCAGGAACCTCAGACCGGTCCTGATATCTCCAGGGCAGCCTGCTGCTGTATCCCTGC	1441
QY	1471	TCCAGGAAAGGCATCTTAAAGAGTCTCGACAGGTGAATCTGGTACTACTCTCTCTC	1530
Db	1442	TCCAGGAAAGGCATCTTAAAGAGTCTCGACAGGTGAATCTGGTACTACTCTCTCTC	1501
QY	1531	CAGAGCCACGAGTCTGGGGAACTCTTAGACGCGAGTGTGTGTGAGTGGGAGCC	1590
Db	1502	CAGAGCCACGAGTCTGGGGAACTCTTAGACGCGAGTGTGTGTGAGTGGGAGCC	1561
QY	1591	CCGTGAGCAGAAAGTCTCCACAGGCTTCAGGGCTCTCTCCACCGCAAGGCAATCTCA	1650
Db	1562	CCGTGAGCAGAAAGTCTCCACAGGCTTCAGGGCTCTCTCCACCGCAAGGCAATCTCA	1621
QY	1651	AACCTAATGSCAAGTCTCCGACAGCCTTAGAGCACTACCCCTAGACACCTTTGGCT	1710
Db	1622	AACCTAATGSCAAGTCTCCGACAGCCTTAGAGCACTACCCCTAGACACCTTTGGCT	1681
QY	1711	CCCTGACCAACTGGCTCTCTCCATCTCGACGCCGGCCAGCCGCCCTCAGGGGCTG	1770
Db	1682	CCCTGACCAACTGGCTCTCTCCATCTCGACGCCGGCCAGCCGCCCTCAGGGGCTG	1741
QY	1771	TGAGTGAGGACAGCATCTGTCTCCGAGTCTTTGACCAATGGACTTGGCTGAACGTC	1830
Db	1742	TGAGTGAGGACAGCATCTGTCTCCGAGTCTTTGACCAATGGACTTGGCTGAACGTC	1801
QY	1831	TTCCCGAAACCCACTGAGGGGTGTGTGTCTGTGGAACCACTGAGGGGCTTTGAGCAGC	1890
Db	1802	TTCCCGAAACCCACTGAGGGGTGTGTGTCTGTGGAACCACTGAGGGGCTTTGAGCAGC	1861
QY	1891	CTCCCTCAGAAGTCTGAAGCGATGTGGCAGGAATCTTGGGGGATAGCTGCTTTTCTC	1950
Db	1862	CTCCCTCAGAAGTCTGAAGCGATGTGGCAGGAATCTTGGGGGATAGCTGCTTTTCTC	1921
QY	1951	TGACAGACTGCCAAGAGGTGACTGCGACCTTACAGAACGCCCTAGGAATCTGCTCAAGC	2010
Db	1922	TGACAGACTGCCAAGAGGTGACTGCGACCTTACAGAACGCCCTAGGAATCTGCTCAAGC	1981
QY	2011	TCAGCTGAGGAAGGAGATGTCCTCTAGTATGGGTAGGCTCTGAGAGGTTTGAGAG	2070
Db	1982	TCAGCTGAGGAAGGAGATGTCCTCTAGTATGGGTAGGCTCTGAGAGGTTTGAGAG	2041

QY	2071	GAAACCTGGGTGGATTCCTCCAGTGAATAGAGTACATCAAGGGCTCTACTCTGCGAGCC	2130
Db	2042	GAACCTGGGTGGATTCCTCCAGTGAATAGAGTACATCAAGGGCTCTACTCTGCGAGCC	2101
QY	2131	TGACTGAACCTGAAAGATGAGAAATTCGATTTGATGTGAAAGAAATGGAAACCTTTGC	2190
Db	2102	TGACTGAACCTGAAAGATGAGAAATTCGATTTGATGTGAAAGAAATGGAAACCTTTGC	2161
QY	2191	TGCCCAGAGTGTATAGTGGGGTGGCTGAGGTGCTTACCTCTTGTGTCATGAGTGTCT	2250
Db	2162	TGCCCAGAGTGTATAGTGGGGTGGCTGAGGTGCTTACCTCTTGTGTCATGAGTGTCT	2221
QY	2251	ACCATGACATTTCCACCCCTGTTCTCTGGCTGCACCTTCCACATAAGTTTCTGTTTCCAT	2310
Db	2222	ACCATGACATTTCCACCCCTGTTCTCTGGCTGCACCTTCCACATAAGTTTCTGTTTCCAT	2281
QY	2311	CAACCAACAGGGTTAGAACCCCTGACTTCTCTGGAGGTAAATGTGTAGTGAATGTCATATT	2370
Db	2282	CAACCAACAGGGTTAGAACCCCTGACTTCTCTGGAGGTAAATGTGTAGTGAATGTCATATT	2341
QY	2371	TAGAGAGGAAACAGCCTCTGGTTTCCATCTCTGCTGCTGTCATCTCAAGACCTGGAA	2430
Db	2342	TAGAGAGGAAACAGCCTCTGGTTTCCATCTCTGCTGCTGTCATCTCAAGACCTGGAA	2401
QY	2431	GACTCCGACCGCTGTTTGAATCTCTCAAGGGGACAGATGCCCTGGACCCCATCTTA	2490
Db	2402	GACTCCGACCGCTGTTTGAATCTCTCAAGGGGACAGATGCCCTGGACCCCATCTTA	2461
QY	2491	GATCTCAGAGACTTGAACCTTGAAGCTTCTTAGTACCAGATGTGGATGGATGCTCTG	2550
Db	2462	GATCTCAGAGACTTGAACCTTGAAGCTTCTTAGTACCAGATGTGGATGGATGCTCTG	2521
QY	2551	TTTCTCAGGCCAAGCGGACCTAGAAATGTGCTGACTTATTTATTTTGTGATCTCACT	2610
Db	2522	TTTCTCAGGCCAAGCGGACCTAGAAATGTGCTGACTTATTTATTTTGTGATCTCACT	2581
QY	2611	TCGTGTTTTTGGTTTTTGTGTTTTGTTTTGTTTTTAAAGTGAATTTTGTGCTGCTT	2670
Db	2582	TCGTGTTTTTGGTTTTTGTGTTTTGTTTTGTTTTTAAAGTGAATTTTGTGCTGCTT	2641
QY	2671	CAATAATGGAATGCTGTGTTCTGGGAACTCCACTGTGCCACTGAAGTTTATGTACAGA	2730
Db	2642	CAATAATGGAATGCTGTGTTCTGGGAACTCCACTGTGCCACTGAAGTTTATGTACAGA	2701
QY	2731	GAAGTATTTGGCAATGATGCTCCCTCTATTCAAGGGGGGTGGGGCGTTTTTCAAAATGAT	2790
Db	2702	GAAGTATTTGGCAATGATGCTCCCTCTATTCAAGGGGGGTGGGGCGTTTTTCAAAATGAT	2761
QY	2791	GTCTTGAGCACTGCTGAGGATGAGTCTCCAGTCCCTTTCACACCAAGGCTGGCCACCTC	2850
Db	2762	GTCTTGAGCACTGCTGAGGATGAGTCTCCAGTCCCTTTCACACCAAGGCTGGCCACCTC	2821
QY	2851	CCTCATCTTTCATCTGTGGCCAAAAAATAAAAAAAAAAAAAAAAAAAAAA 2899	
Db	2822	CCTCATCTTTCATCTGTGGCCAAAAAATAAAAAAAAAAAAAAAAAAAAAA 2870	

## RESULT 2

AK033672

LOCUS

DEFINITION

AK033672

VERSION

AK033672.1

GI:26329364

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus

ORGANISM

Mus musculus (house mouse)

REFERENCE

1

AUTHORS

Carninci, P. and Hayashizaki, Y.

AK033672 3098 bp mRNA linear HTC 03-APR-2004  
Mus musculus adult male cecum cDNA, RIKEN full-length enriched  
library, clone:9130215K18 product:weakly similar to PROBABLE  
SERINE/THROMBIN-PROTEIN KINASE KIAA0537 (EC 2.7.1.1.) (Homo  
sapiens), full insert sequence.

AK033672

AK033672.1

GI:26329364

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus

ORGANISM

Mus musculus (house mouse)

REFERENCE

1

AUTHORS

Carninci, P. and Hayashizaki, Y.



Qy	631	GCACCAAGACGGGATCGTTACCGAGATCTCAAGCTGGAAAAACATCCTTCTAGATGCCA	690
Db	601	GCACCAAGACGGGATCGTTACCGAGATCTCAAGCTGGAAAAACATCCTTCTAGATGCCA	660
Qy	691	ATGGAACATCAAGATTGCTGACCTTTGGGCTCTCCAACTGTACCAAAAGGCAAGTTCC	750
Db	661	ATGGAACATCAAGATTGCTGACCTTTGGGCTCTCCAACTGTACCAAAAGGCAAGTTCC	720
Qy	751	TCCAGAGCTTCTGTGGGAGCCCTCTACGCGCTCGCTGAGATAGTCAACGGGAAGCCCT	810
Db	721	TCCAGAGCTTCTGTGGGAGCCCTCTACGCGCTCGCTGAGATAGTCAACGGGAAGCCCT	780
Qy	811	ATGTGGGCCAGAGGTTGACAGCTGGTCTCTGGGCGTTCTCTGTATACATCCTGTGTGATG	870
Db	781	ATGTGGGCCAGAGGTTGACAGCTGGTCTCTGGGCGTTCTCTGTATACATCCTGTGTGATG	840
Qy	871	GCACCATGCCCTTTGACGGGAGGATCATAAAAACATCTGGTGAAGCAATCAGTAAACGGG	930
Db	841	GCACCATGCCCTTTGACGGGAGGATCATAAAAACATCTGGTGAAGCAATCAGTAAACGGG	900
Qy	931	CTTACCGTGAAGCGCCCAAGCCGTCAGATGCTGTGGCTGATCCGGTGGCTGTTAATGG	990
Db	901	CTTACCGTGAAGCGCCCAAGCCGTCAGATGCTGTGGCTGATCCGGTGGCTGTTAATGG	960
Qy	991	TGAACCCCAACCGTCCGGGCCACACCTGGAGGATGTAGCCAGTCAATTGGTGGGTCAACTGGG	1050
Db	961	TGAACCCCAACCGTCCGGGCCACACCTGGAGGATGTAGCCAGTCAATTGGTGGGTCAACTGGG	1020
Qy	1051	GTACACCAACCGGAGTCCGGGAAACAGAAAGCCCTGCGTGAAGGTTGGCAACCTAGTGGTG	1110
Db	1021	GTACACCAACCGGAGTCCGGGAAACAGAAAGCCCTGCGTGAAGGTTGGCAACCTAGTGGTG	1080
Qy	1111	ACTTTGGCCGGGCTCCATGGCGGACTGTTAGTCTGCTCTCGCGCCCTCTCTGGAGA	1170
Db	1081	ACTTTGGCCGGGCTCCATGGCGGACTGTTAGTCTGCTCTCGCGCCCTCTCTGGAGA	1140
Qy	1171	ATGAGCAAGAGTGTGAGCTTCTTCAAGCAGACGCTGCGGGAGGTGGAAGCACTGTAC	1230
Db	1141	ATGAGCAAGAGTGTGAGCTTCTTCAAGCAGACGCTGCGGGAGGTGGAAGCACTGTAC	1200
Qy	1231	CTGGGCTGGAGCGGCAAACTTCTTTAAGAAAGTCCGAAAGGAGAAATGAATGGCTCAA	1290
Db	1201	CTGGGCTGGAGCGGCAAACTTCTTTAAGAAAGTCCGAAAGGAGAAATGAATGGCTCAA	1260
Qy	1291	ATCTGCAAGTGAACCGGCTGAGGATACCTTCTCGCCCTGCGCAAGCAGCCTTAAGC	1350
Db	1261	ATCTGCAAGTGAACCGGCTGAGGATACCTTCTCGCCCTGCGCAAGCAGCCTTAAGC	1320
Qy	1351	TTCCGAAAGGCAATCTCAAGAAAGTCCCTACCTCGTCAAGGGAGGTACAGGAGACC	1410
Db	1321	TTCCGAAAGGCAATCTCAAGAAAGTCCCTACCTCGTCAAGGGAGGTACAGGAGACC	1380
Qy	1411	CTCAGGAACCTCAGACCGGCTGCTGATCTCAAGGAGCCTGCTCCCTGCTGTATCCCTGC	1470
Db	1381	CTCAGGAACCTCAGACCGGCTGCTGATCTCAAGGAGCCTGCTCCCTGCTGTATCCCTGC	1440
Qy	1471	TCCCAAGGAAGGCAATCTTAAAGAGTCTCGACAGCGTGAATCTGGTTACTCTCTCTC	1530
Db	1441	TCCCAAGGAAGGCAATCTTAAAGAGTCTCGACAGCGTGAATCTGGTTACTCTCTCTC	1500
Qy	1531	CAGAGCCAGCGAGTCTGGGGAACCTTTAGACGCGAGTGTGTTGTGAGTGGGAGCC	1590
Db	1501	CAGAGCCAGCGAGTCTGGGGAACCTTTAGACGCGAGTGTGTTGTGAGTGGGAGCC	1560
Qy	1591	CCGTGGAGCAAGTCTCCACAGGCTTCAGGGCTCTCTCCACCGCAAGGCAATCTCA	1650
Db	1561	CCGTGGAGCAAGTCTCCACAGGCTTCAGGGCTCTCTCCACCGCAAGGCAATCTCA	1620
Qy	1651	AACTCAATGGCAAGTCTCCGCGACAGCTTTAGAGGCACTACCCCTAGACACCTTTGGCT	1710
Db	1621	AACTCAATGGCAAGTCTCCGCGACAGCTTTAGAGGCACTACCCCTAGACACCTTTGGCT	1680

Qy	1711	CCCTGGACCAACTGGCTCTCTCCATCTCTGACGCGCGGCCAGCGCCCCCTCAGGGGCTG	1770
Db	1681	CCCTGGACCAACTGGCTCTCTCCATCTCTGACGCGCGGCCAGCGCCCCCTCAGGGGCTG	1740
Qy	1771	TGAGTGAGGACAGCATCTCTCCGAGTCTCTTGCAGTCTTTCACCAATTTGGACTTGCCTGAACGTC	1830
Db	1741	TGAGTGAGGACAGCATCTCTCCGAGTCTCTTGCAGTCTTTCACCAATTTGGACTTGCCTGAACGTC	1800
Qy	1831	TTCCCGAAACCCCACTGAGGGGCTGTGTCTGTGCAACACCTGAGGGGCTTTGAGCAGC	1890
Db	1801	TTCCCGAAACCCCACTGAGGGGCTGTGTCTGTGCAACACCTGAGGGGCTTTGAGCAGC	1860
Qy	1891	CTCCCTCAGAAGTCTGAAGCGATGTGGCAGGAATCTTTGGGGATAGTGTCTTTTCTC	1950
Db	1861	CTCCCTCAGAAGTCTGAAGCGATGTGGCAGGAATCTTTGGGGATAGTGTCTTTTCTC	1920
Qy	1951	TGACACACTGCCAAGAGGTGACTGACGCTACAGACCAAGCCCTAGGAATCTGCTCAAGC	2010
Db	1921	TGACACACTGCCAAGAGGTGACTGACGCTACAGACCAAGCCCTAGGAATCTGCTCAAGC	1980
Qy	2011	TCACTGAGGAAGGAGATGGTCCCTAGTATGGGGTAGGCTCTGAGAGGGTTTGACAG	2070
Db	1981	TCACTGAGGAAGGAGATGGTCCCTAGTATGGGGTAGGCTCTGAGAGGGTTTGACAG	2040
Qy	2071	GAAACCTGGGTCCGATTTCTCCAGTGAATAGAGTACATCAAGGGCTCTACGTCTGCAGCC	2130
Db	2041	GAAACCTGGGTCCGATTTCTCCAGTGAATAGAGTACATCAAGGGCTCTACGTCTGCAGCC	2100
Qy	2131	TGACTGAACTCTGAAGATGAGAAATCGGATTGATGTGGAAAGGAATGGGAACCTTGC	2190
Db	2101	TGACTGAACTCTGAAGATGAGAAATCGGATTGATGTGGAAAGGAATGGGAACCTTGC	2160
Qy	2191	TGCCCAGAGTGTATAGTGGGGTGGCTGAAGGTGCTACCTCTTTGTGTCATGAGTGC	2250
Db	2161	TGCCCAGAGTGTATAGTGGGGTGGCTGAAGGTGCTACCTCTTTGTGTCATGAGTGC	2220
Qy	2251	ACCATGACATTTCCACACCTCTCTGCTGCTCACATAAGTTTCTGTCTTCCAT	2310
Db	2221	ACCATGACATTTCCACACCTCTCTGCTGCTCACATAAGTTTCTGTCTTCCAT	2280
Qy	2311	CAACCAAGGTTAGAACCTGACTTCTGGAGGTAAATGTGTAGTACTGCGATTTAT	2370
Db	2281	CAACCAAGGTTAGAACCTGACTTCTGGAGGTAAATGTGTAGTACTGCGATTTAT	2340
Qy	2371	TAGAGAGGAAACAGCTCTGGTTTCCATCTCTGCTGTGTCATCTCAAGACCTGGAA	2430
Db	2341	TAGAGAGGAAACAGCTCTGGTTTCCATCTCTGCTGTGTCATCTCAAGACCTGGAA	2400
Qy	2431	GACTCGGACCGCTGTTTGACTTCTCAAGGGGACAGATGCCCTGGACCCCATCTTA	2490
Db	2401	GACTCGGACCGCTGTTTGACTTCTCAAGGGGACAGATGCCCTGGACCCCATCTTA	2460
Qy	2491	GATCTCAGAGACTTGAACCTTGAAGCTTCTTAGTACCAGATGTGGATGGATGCTCTG	2550
Db	2461	GATCTCAGAGACTTGAACCTTGAAGCTTCTTAGTACCAGATGTGGATGGATGCTCTG	2520
Qy	2551	TTTCTCAGGCCAACGGGACCTGAAGTGTCTGACTTATTTATTTTGTGTTTCTCACT	2610
Db	2521	TTTCTCAGGCCAACGGGACCTGAAGTGTCTGACTTATTTATTTTGTGTTTCTCACT	2580
Qy	2611	CTGTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTTAAAGTGAATTTTGTGCTTTT	2670
Db	2581	CTGTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTTAAAGTGAATTTTGTGCTTTT	2640
Qy	2671	CAATAATGGAATGTCTGTGTTTGGGAACTCCACTGTGCGCACTGAAGTTTATGTACAGA	2730
Db	2641	CAATAATGGAATGTCTGTGTTTGGGAACTCCACTGTGCGCACTGAAGTTTATGTACAGA	2700
Qy	2731	GAGTATTTGGCAATGATGCTCCCTCTATTCAAGGGGGGTGGGGCGTTTTTCAATGTAT	2790
Db	2701	GAGTATTTGGCAATGATGCTCCCTCTATTCAAGGGGGGTGGGGCGTTTTTCAATGTAT	2760
Qy	2791	GTCTTGAGCACTGTCTGGATTGAGTCTCCAGTCCCTTACACCCCAAGGCTGGCCACCTC	2850

```

Db      2761  GTCTGAGCACTGCTGATTTAGTCTCCAGTCCCTTACACCAAGCGTGGCCACCTC 2820
|||||
Qy      2851  CCTCATCTTCATCTGTGGCCAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 2899
|||||
Db      2821  CCTCATCTTCATCTGTGGCCAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 2869
|||||

RESULT 3
LOCUS   AK034082
DEFINITION
Mus musculus adult male diencephalon cDNA, RIKEN full-length
enriched library, clone:9330154N24 product:weakly similar to
PROBABLE SERINE/THREONINE-PROTEIN KINASE KIAA0537 (EC 2.7.1.-)
[ Homo sapiens ], full insert sequence.
ACCESSION
VERSION AK034082.1 GI:26329672
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 92279253
PUBMED 10349636
REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE
4 The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE
6 (bases 1 to 2869)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Nomura, R., Ohno, M., Oheato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

```

```

Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
FEATURES
Location/Qualifiers
1..2869
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:9330154N24"
/db_xref="taxon:10090"
/clone="9330154N24"
/sex="male"
/tissue_type="dienecephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
98..2017
/notes="unnamed protein product; putative
weakly similar to PROBABLE SERINE/THREONINE-PROTEIN KINASE
KIAA0537 (EC 2.7.1.-) [Homo sapiens] (SWISSPROT|060285,
evidence: FASTY, 56.8%ID, 98.6%length, match=1849)."
/codon_start=1
/protein_id="BAC28575.1"
/db_xref="GI:26329673"
/translation="MESGALLQPSQAPSASALASBPRLADGLIKSPKPLMKQAV
KHHHKNHRYELETGKTYGKVKARESSGLVAIKSIKDKIKDDODLHIR
REIEMSSLNHPHIIAHHVGRSLVTVENSKIVVMEYASRGDIYISRPRLS
EDARHFFQIVSALHYCHQNGIVHRLDKLNLDDANGNIKIADGLNLNKKFL
QTFCEPSPASPEIVNGPYGVDSVLSGLVLLYLVHGTVPDQDHTLVKQISN
GYRREPSPDAGLIRWLLNPTTRATLEDVASHWVNWGYTTGVGQELKRSKH
PSGDFRASMADLRSSPRLLENGAKVCSFKQHPVGGSTVGLRGHSLKRSKH
NMAQNLQGDPAEDTSRRPCKSLKPKGILKKSSTSGVEQDPQELRPVDPDTGQ
PVPVSLLRKGLIKSRQESGYSSPEPSESGELLADSDVPSDGPVEQKSPQASG
LLHHRKGLIKKFSRTALEGTTFTFGLDQLASSHPAARPSRPSGVAESDSIUS
ESFDQDLPLPETLPETLPGCVSDNLRGLQPPSEGLKRWQESLGDSCFSLTDCQEV
TAAVROALGICSKLS"
ORIGIN
Query Match 96.6%; Score 2804.4; DB 3; Length 2869;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 2839; Conservative 0; Mismatches 1; Indels 24; Gaps 1;
Qy 31 GCCTGCTCGGGTGGCTGTGACCTCTGAGCCCGCGGCTCAGCGCGGCTGCTGCTGCTG 90
|||||
Db 6 GCCTGCTCGGGTGGCTGTGACCTCTGAGCCCGCGGCTCAGCGCGGCTGCTGCTGCTG 65
|||||
Qy 91 CCGGACCCACTCCACTCGCGGTCCCGGCACCATGAGTGGTGGGCTTACTTACTCAGCGCC 150
|||||
Db 66 CCGGACCCACTCCACTCGCGGTCCCGGCACCATGAGTGGGCTTACTTACTCAGCGCC 125
|||||
Qy 151 CGAGCCAGGCTCCCTCGGCTCGGCTCGGCTCGGAGAGCGCCCGCGCGCTGGCGGACG 210
|||||
Db 126 CGAGCCAGGCTCCCTCGGCTCGGCTCGGCTCGGAGAGCGCCCGCGCGCTGGCGGACG 185
|||||
Qy 211 GGCTCATCAAGTGGCTTAAACCTCTCATGAAGAAGCAGCGGCTGAAGCGGCACCATCACA 270
|||||
Db 186 GGCTCATCAAGTGGCTTAAACCTCTCATGAAGAAGCAGCGGCTGAAGCGGCACCATCACA 245
|||||
Qy 271 AACACAACCTGGGCGACCGCTAGAGTCTCTGGAGACGCTGGGCGAAGGGACCTACGGGA 330
|||||
Db 246 AACACAACCTGGGCGACCGCTAGAGTCTCTGGAGACGCTGGGCGAAGGGACCTACGGGA 305
|||||

```

QY 331 AGGTGAAGAGGACGAGAGAGCTCGGGGGTCTGGTGGCCATCAAGTCCATCAGGAAAG 390  
DB 306 AGGTGAAGAGGACGAGAGAGCTCGGGGGTCTGGTGGCCATCAAGTCCATCAGGAAAG 365  
QY 391 ACATAATCAAGATGAGCAGGATCTGTGCACATACGAGGGGAGATTGAGATCATGTCTT 450  
DB 366 ACAAAATCAAGATGAGCAGGATCTGTGCACATACGAGGGGAGATTGAGATCATGTCTT 425  
QY 451 CACTCAACCCACCCACATCATTCGCCATCCATGAA-----G 486  
DB 426 CACTCAACCCACCCACATCATTCGCCATCCATGAGTGGGAGATCAAGCTTAGTGACTG 485  
QY 487 TGTGTGAGATPAGCAGCAAGATTGTGATTGTCTGAGATATGCCAGCCGAGGCGATCTGT 546  
DB 486 TGTGTGAGATPAGCAGCAAGATTGTGATTGTCTGAGATATGCCAGCCGAGGCGATCTGT 545  
QY 547 ATGATTACATCAGTAGGCGGCCACGCTGTGATGAGTGGGAGCGGAGCCAGGATTTCTCCGAC 606  
DB 546 ATGATTACATCAGTAGGCGGCCACGCTGTGATGAGTGGGAGCGGAGGATTTCTTCGAC 605  
QY 607 AGATCTGTCTGCCCTGCATCTACTGCCACAGAAACGGGATCGTTACCGAGATCTCAAGC 666  
DB 606 AGATCTGTCTGCCCTGCATCTACTGCCACAGAAACGGGATCGTTACCGAGATCTCAAGC 665  
QY 667 TGGAAAAACATCTTCTAGATGCCAATGGAACATCAAGATTGCTGACTTTGGCCTCTCCA 726  
DB 666 TGGAAAAACATCTTCTAGATGCCAATGGAACATCAAGATTGCTGACTTTGGCCTCTCCA 725  
QY 727 ACCTGTACCAAAAGGCAAGTTCTCCAGAGTTCTGTGGAGCCCTCTCTAGCCCTGCG 786  
DB 726 ACCTGTACCAAAAGGCAAGTTCTCCAGAGTTCTGTGGAGCCCTCTCTAGCCCTGCG 785  
QY 787 CTGAGATAGTCAACGGGAAGCCCTATCTGTGGGCCAGAGGTGGACAGTCTCTCTGGCG 846  
DB 786 CTGAGATAGTCAACGGGAAGCCCTATCTGTGGGCCAGAGGTGGACAGTCTCTCTGGCG 845  
QY 847 TTCTCTGTATCATCTCTGGTGCATGGCACCATGCCCCTTTGACGGGAGGATCATAAACAC 906  
DB 846 TTCTCTGTATCATCTCTGGTGCATGGCACCATGCCCCTTTGACGGGAGGATCATAAACAC 905  
QY 907 TGGTGAAGCAAAATCAGTAAACGGGCTTACCGTAGCGCCGCCAAGCCGCTCGATCTGTG 966  
DB 906 TGGTGAAGCAAAATCAGTAAACGGGCTTACCGTAGCGCCGCCAAGCCGCTCGATCTGTG 965  
QY 967 GCCTGATCCGGTGGCTGTTAATGTTGAACCCACCGTCCGGGCCACACTGGAGGATGTAG 1026  
DB 966 GCCTGATCCGGTGGCTGTTAATGTTGAACCCACCGTCCGGGCCACACTGGAGGATGTAG 1025  
QY 1027 CCAGTCAATTGGTGGGTCAACTGGGGTTACACCCGGAGTCCGGGAACAGGAAAGCCCTGC 1086  
DB 1026 CCAGTCAATTGGTGGGTCAACTGGGGTTACACCCGGAGTCCGGGAACAGGAAAGCCCTGC 1085  
QY 1087 GTGAGGTGGGACCCCTAGTGTGACTTTGGCGGGCTCCATGGGGGACTGTGTACGTC 1146  
DB 1086 GTGAGGTGGGACCCCTAGTGTGACTTTGGCGGGCTCCATGGGGGACTGTGTACGTC 1145  
QY 1147 GCTCTCGCCGCCCTCTCTGGAGATGGAGCCAAAGTGTGCAGCTTCTCAACGACGACG 1206  
DB 1146 GCTCTCGCCGCCCTCTCTGGAGATGGAGCCAAAGTGTGCAGCTTCTCAACGACGACG 1205  
QY 1207 TGCCGGGAGGTGAAGACATGTACTCTGGGTGGAGCGGCAACATTTCTTTAAGAAGTCCC 1266  
DB 1206 TGCCGGGAGGTGAAGACATGTACTCTGGGTGGAGCGGCAACATTTCTTTAAGAAGTCCC 1265  
QY 1267 GAAAGGAGATGATAGTGTCAAAATCTGAAAGGTGACCGGCTGAGGATACCTCTTCTC 1326  
DB 1266 GAAAGGAGATGATAGTGTCAAAATCTGAAAGGTGACCGGCTGAGGATACCTCTTCTC 1325  
QY 1327 GCCCTGCAAGAGCAGCCTTAAGCTTCCGAAAGGCAATCTCAAGAAAGTCTCTACCT 1386  
DB 1326 GCCCTGCAAGAGCAGCCTTAAGCTTCCGAAAGGCAATCTCAAGAAAGTCTCTACCT 1385

QY 1387 CGTCAGGGAGGTACAGGAGGACCCCTCAGGAATCAGACCGGTGCCTGATACTCCAGGC 1446  
DB 1386 CGTCAGGGAGGTACAGGAGGACCCCTCAGGAATCAGACCGGTGCCTGATACTCCAGGC 1445  
QY 1447 AGCTGTCCCTGTGTATCTCTCTCCAAAGAAAGCATCTTTAAGAGTCTCGACAGC 1506  
DB 1446 AGCTGTCCCTGTGTATCTCTCTCCAAAGAAAGCATCTTTAAGAGTCTCGACAGC 1505  
QY 1507 GTGAATCTGTGTACTACTCTCTCCAGAGCCAGGAGTCTGGGAACTCTTAGACCCA 1566  
DB 1506 GTGAATCTGTGTACTACTCTCTCCAGAGCCAGGAGTCTGGGAACTCTTAGACCCA 1565  
QY 1567 GTGATGTGTTTGTAGTGGGACCCCGTGGAGCAGAGTCTCCACAGGCTTCAGGGCTCC 1626  
DB 1566 GTGATGTGTTTGTAGTGGGACCCCGTGGAGCAGAGTCTCCACAGGCTTCAGGGCTCC 1625  
QY 1627 TCCTCCACCCCAAGGGGATTTCTCAAACTCAATGGCAAGTCTCCCGCACAGCCTTAGAAG 1686  
DB 1626 TCCTCCACCCCAAGGGGATTTCTCAAACTCAATGGCAAGTCTCCCGCACAGCCTTAGAAG 1685  
QY 1687 GCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGGCTCTCCCATCTCTGACGCC 1746  
DB 1686 GCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGGCTCTCCCATCTCTGACGCC 1745  
QY 1747 GGCCACAGCCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTCTCTCCAGTCTCTTTG 1806  
DB 1746 GGCCACAGCCGCCCTCAGGGGCTGTGAGTGAGGACAGCATCTCTCTCCAGTCTCTTTG 1805  
QY 1807 ACCAATTGGAATTGCTGTGAACGCTTCCCGAAACCCCACTGAGGGGCTGTGTCTGTGG 1866  
DB 1806 ACCAATTGGAATTGCTGTGAACGCTTCCCGAAACCCCACTGAGGGGCTGTGTCTGTGG 1865  
QY 1867 ACAACCTGAGGGGCTTGAGCGCTCCCTCAGAAGGTCTGAAAGCGATGTGGCAGGAAT 1926  
DB 1866 ACAACCTGAGGGGCTTGAGCGCTCCCTCAGAAGGTCTGAAAGCGATGTGGCAGGAAT 1925  
QY 1927 CCTTGGGGATAGCTGCTTTCTGCACAGACTGCCAAGAGGTGACTGACAGCTTACAGAC 1986  
DB 1926 CCTTGGGGATAGCTGCTTTCTGCACAGACTGCCAAGAGGTGACTGACAGCTTACAGAC 1985  
QY 1987 AGCCCTTAGGAATCTGCTCAAAAGCTCAGCTGAGGAGGAGATGGTCCCTAGTATGGGG 2046  
DB 1986 AGCCCTTAGGAATCTGCTCAAAAGCTCAGCTGAGGAGGAGATGGTCCCTAGTATGGGG 2045  
QY 2047 TAGGCTCTCAGAGGGTTTCAGAGGAACCTCGGTGGATTCCTCAGTCAANTAGAGTAC 2106  
DB 2046 TAGGCTCTCAGAGGGTTTCAGAGGAACCTCGGTGGATTCCTCAGTCAANTAGAGTAC 2105  
QY 2107 ATCAAGGGCTTACGCTCTGAGGCTGAACTGAACTGAAAGATGAGAGAAATCGCATTTGAT 2166  
DB 2106 ATCAAGGGCTTACGCTCTGAGGCTGAACTGAACTGAAAGATGAGAGAAATCGCATTTGAT 2165  
QY 2167 GTGGAAGGAATGGGAACCCCTGCTGCCGAGTGTATAGTGGGGTGGCTGAAAGGTGCC 2226  
DB 2166 GTGGAAGGAATGGGAACCCCTGCTGCCGAGTGTATAGTGGGGTGGCTGAAAGGTGCC 2225  
QY 2227 TACCTCTTGTGCCATGAGTGCACCATGATTTCCACCCCTGCTCTCTGGCTGCAC 2286  
DB 2226 TACCTCTTGTGCCATGAGTGCACCATGATTTCCACCCCTGCTCTCTGGCTGCAC 2285  
QY 2287 CTTTACATTAAGTTTCTGTTTCCATCAACCAAGGGTTAGAACCTTGACTTCTCTGGGAGG 2346  
DB 2286 CTTTACATTAAGTTTCTGTTTCCATCAACCAAGGGTTAGAACCTTGACTTCTCTGGGAGG 2345  
QY 2347 TAATGTAGTGAATGCCATTTATTAGAGAGGAAACAGGCTCTGGTTTCCATCTCTGCTG 2406  
DB 2346 TAATGTAGTGAATGCCATTTATTAGAGAGGAAACAGGCTCTGGTTTCCATCTCTGCTG 2405  
QY 2407 CTGTGATCTCAAGAGACCTGGGAGACTCGGACCGCTGTTTGAATCTCATCTCAAGGGAC 2466  
DB 2406 CTGTGATCTCAAGAGACCTGGGAGACTCGGACCGCTGTTTGAATCTCATCTCAAGGGAC 2465  
QY 2467 CAGATGCCCTGGACCCCATCTTAGATCTCAGAGACTTGAACTTGAAAGCTGTTCTCTAGT 2526

```
Db 2466 CAGATGCCCTGGACCCCACTTAGATCTCAGAGACTTGAACCTTGAAGCTGTTCTTAGT 2525
Qy 2527 ACCAGATGGATGGATGCTCTGTTTCTCAGGCAACAGGACCTAGAATGTGCTGACTT 2586
Db 2526 ACCAGATGGATGGATGCTCTGTTTCTCAGGCAACAGGACCTAGAATGTGCTGACTT 2585
Qy 2587 ATTATATTTTTTGTGATCTCAGTCTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 2646
Db 2586 ATTATATTTTTTGTGATCTCAGTCTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 2645
Qy 2647 TTTTAAAGTGAATTTTCTGCTTTCAATAATGTGAATGCTGTCTGTTTGTGTTTGTGTTTGTG 2706
Db 2646 TTTTAAAGTGAATTTTCTGCTTTCAATAATGTGAATGCTGTCTGTTTGTGTTTGTGTTTGTG 2705
Qy 2707 GTGCCACTGAAGTTTATGTACAGAGAAATTTGGCAATGATGTCCTCTATTCAAGGGG 2766
Db 2706 GTGCCACTGAAGTTTATGTACAGAGAAATTTGGCAATGATGTCCTCTATTCAAGGGG 2765
Qy 2767 GGTGGGGCGGTTTTTCAAAATGATGCTCTGAGCACTGTCTGGATTGAGTCTCCAGTCCCT 2826
Db 2766 GGTGGGGCGGTTTTTCAAAATGATGCTCTGAGCACTGTCTGGATTGAGTCTCCAGTCCCT 2825
Qy 2827 TCACACCAAGGCTGGCCACCCTCCCTCATCTTCATCTCTGTGGCC 2870
Db 2826 TCACACCAAGGCTGGCCACCCTCCCTCATCTTCATCTGTGGCC 2869

RESULT 4
AY399041
LOCUS AY399041 1896 bp DNA linear GSS 12-DEC-2003
DEFINITION Mus musculus HCM0091 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY399041
VERSION AY399041.1 GI:39755030
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1896)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
location/Qualifiers
1..1896
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<!.>1896
/locus_tag="HCM0091"
gene
ORIGIN
Query Match 61.2%; Score 1776; DB 9; Length 1896;
Best Local Similarity 93.7%; Pred No. 0;
Matches 1776; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
Qy 123 ATGGAGTCGGTGGCCTTACTCAGCGCCGAGCCAGGCTCCCTCGGCCTCGCCTCGGCC 182
```

```
1 ATGGAGTCGGTGGCCTTACTCCAGCGCCGAGCCAGGCTCCCTCGGCCTCGCCTCGGCC 60
183 TGGAGAGCGCCCGCGCGCTGCGGACGCGGCTCATCAAGTCGCTTAAACCTCTGATGAG 242
61 TCGGAGAGCGCCCGCGCGCTGCGGACGCGGCTCATCAAGTCGCTTAAACCTCTGATGAG 120
243 AAGCAGCGCGTGAAGCGGCAACCATCAAAACAACCTCGCGCACCGCTACAGATTCCTG 302
121 AAGCAGCGGTGAAGCGGCAACCATCAAAACAACCTCGCGCACCGCTACAGATTCCTG 180
303 GAGACGCTCGGCAAGGGCACTTA CGGGAAGGTGAAGAAGGCA CGAGAGAGCTCGGGGCGT 362
181 GAGACGCTCGGCAAGGGCACTTA CGGGAAGGTGAAGAAGGCA CGAGAGAGCTCGGGGCGT 240
363 CTGGTGGCCATCAAGTCCCATCAGGAAGACAAAATCAAAAGATGAGCAGATCTCTGAC 422
241 CTGGTGGCCATCAAGTCCCATCAGGAAGACAAAATCAAAAGATGAGCAGATCTCTGAC 300
423 ATACGAGGAGGAGATTGAGATCATGTCTTCACTCAACACCCCACTATTGTCATCCAT 482
301 ATACGAGGAGGAGATTGAGATCATGTCTTCACTCAACACCCCACTATTGTCATCCAT 360
483 GAAGTCTTTGAGAATAGCAGCAAGATTGTGATGATGATGATGATGATGATGATGATGAT 542
361 GAAGTCTTTGAGAATAGCAGCAAGATTGTGATGATGATGATGATGATGATGATGATGAT 420
543 CTGTATGATTTACATCAGTGCAGGCGCCACGCTGAGTGAGCGGACCCAGGCAATTTCTTC 602
421 CTGTATGATTTACATCAGTGCAGGCGCCACGCTGAGTGAGCGGACCCAGGCAATTTCTTC 480
603 CGACAGATCGTGTCTGCCCTGCACTACTGCCACAGAACGGGATCGTTTCAACGAGATCTC 662
481 CGACAGATCGTGTCTGCCCTGCACTACTGCCACAGAACGGGATCGTTTCAACGAGATCTC 540
663 AAGCTGGAAAAACATCTCTTAGATGCCAATGGAACATCAAGATTGCTGACTTTGGCCTC 722
541 AAGCTGGAAAAACATCTCTTAGATGCCAATGGAACATCAAGATTGCTGACTTTGGCCTC 600
723 TCCAACTGTACCACAAAGCAAGTCTCTCCAGACGCTTCTGTGGGAGCCCTCTCTACGCC 782
601 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 660
783 TCGCCTGAGATAGTCAACGGGAAGCCCTATGTGGGCGCCAGAGGTGCAAGCTGCTCTG 842
661 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 720
843 GCGCTTCTCTGTACATCTGTGTGATGCGACCATGCGCTTTTGAACGGGAGGATCATAA 902
721 GCGCTTCTCTGTACATCTGTGTGATGCGACCATGCGCTTTTGAACGGGAGGATCATAA 780
903 ACACCTGTGAAGCAAAATCAGTAACGGGCTTACCGTGAGCGCGCCCAAGCGCTCGGATGCC 962
781 ACACCTGTGAAGCAAAATCAGTAACGGGCTTACCGTGAGCGCGCCCAAGCGCTCGGATGCC 840
963 TGTGGCCTGATCCGGTGGCTGTTTAATGTTGAACCCCAACCGCTCGGGCACACTGGAGGAT 1022
841 TGTGGCCTGATCCGGTGGCTGTTTAATGTTGAACCCCAACCGCTCGGGCACACTGGAGGAT 900
1023 GTAGCCAGTCAATTTGGTGGGTCAACTGGGTTTACACCAACCGGAGTCCGGGAACAGGAGCC 1082
901 GTAGCCAGTCAATTTGGTGGGTCAACTGGGTTTACACCAACCGGAGTCCGGGAACAGGAGCC 960
1083 CTGCGTGAGGCTGGGCAACCTTAGTGTGACTTTTGGCGGGCCCTCCATGGCGGAGTGTGTTA 1142
961 CTGCGTGAGGCTGGGCAACCTTAGTGTGACTTTTGGCGGGCCCTCCATGGCGGAGTGTGTTA 1020
1143 CGTGTGCTCTCGCGCCCTCTCTGAGAAATGAGAGCCAGAGTGTGAGCTTCTTCAAGCAG 1202
1021 CGTGTGCTCTCGCGCCCTCTCTGAGAAATGAGAGCCAGAGTGTGAGCTTCTTCAAGCAG 1080
1203 CAGTGTGCGGGAGGTGGAGCACTGTACCTGGGCTGAGCGGCAACATCTCTTAAAGAAG 1262
```



```
Db 1081 CAGTGGGAGGTGAAGCACTGTACCTGGGCTGAGCGGCAACATTCCTCTTAAGAAG 1140
Qy 1263 TCCGAAAGGAGATGACATGGCTCAAAATCTCAAGGTGACCCGGCTGAGGATACCTCT 1322
Db 1141 TCCGAAAGGAGATGACATGGCTCAAAATCTCAAGGTGACCCGGCTGAGGATACCTCT 1200
Qy 1323 TCTGCGCTGCAAGAGCAGCTTAAGTTCGAAAGGCAATTCCTCAAGAAAGTCTCTCT 1382
Db 1201 TCTGCGCTGCAAGAGCAGCTTAAGTTCGAAAGGCAATTCCTCAAGAAAGTCTCTCT 1260
Qy 1383 ACCTCGTCAGGGAGGTACAGGAGACCTCAGGAATCTCAGACGGTGCCTGATCTCCA 1442
Db 1261 ACCTCGTCAGGGAGGTACAGGAGACCTCAGGAATCTCAGACGGTGCCTGATCTCCA 1320
Qy 1443 GGGCAGCGTCTCCCTGCTGTATCTCCTCCCAAGGAAAGGCATCTTAAGAACTCTCGA 1502
Db 1321 GGGCAGCGTCTCCCTGCTGTATCTCCTCCCAAGGAAAGGCATCTTAAGAACTCTCGA 1380
Qy 1503 CAGCGTGAATCTGTTACTACTCTCTCAGAGCCCAAGCGAGTCTGGGGAACCTTTAGAC 1562
Db 1381 CAGCGTGAATCTGTTACTACTCTCTCAGAGCCCAAGCGAGTCTGGGGAACCTTTAGAC 1440
Qy 1563 GCCAGTGAATGTTTGTAGTGGGACCCCGTGAGCAGAGTCTCCACAGGCTTCAGGG 1622
Db 1441 GCCAGTGAATGTTTGTAGTGGGACCCCGTGAGCAGAGTCTCCACAGGCTTCAGGG 1500
Qy 1623 CTCCTCTCCACCGCAAGGGCAATTCCTCAAACTCAATGGCAAGTCTCCCGCACAGCCTTA 1682
Db 1501 CTCCTCTCCACCGCAAGGGCAATTCCTCAAACTCAATGGCAAGTCTCCCGCACAGCCTTA 1560
Qy 1683 GAAGGCACTACCCCTAGCAGCTTTGGCTTCCTTGAGCAAACTGGGCTCTCCCTCACTGCA 1742
Db 1561 GAAGGCACTACCCCTAGCAGCTTTGGCTTCCTTGAGCAAACTGGGCTCTCCCTCACTGCA 1620
Qy 1743 GCCGGCCAGCGCCCTCAGGGCTGTGAGTGAGCAGCAGCATCTCTCTCGAGTCC 1802
Db 1621 GCCGGCCAGCGCCCTCAGGGCTGTGAGTGAGCAGCAGCATCTCTCTCGAGTCC 1680
Qy 1803 TTTGACCAATTTGACTTGGCTGACGCTTTCCCGAAACCCCACTGAGGGGCTGTGTCT 1862
Db 1681 TTTGACCAATTTGACTTGGCTGACGCTTTCCCGAAACCCCACTGAGGGGCTGTGTCT 1740
Qy 1863 GTGCAAACTTGAGGGGCTTGAGCAGCCTCCCTCAGAAAGTCTGAAGCGATGGTGACAG 1922
Db 1741 GTGCAAACTTGAGGGGCTTGAGCAGCCTCCCTCAGAAAGTCTGAAGCGATGGTGACAG 1800
Qy 1923 GAATCTTTGGGGATAGTCTCTTTCTCAGACAGCTGCCAAGAGTGTGATGAGCCTAC 1982
Db 1801 GAATCTTTGGGGATAGTCTCTTTCTCAGACAGCTGCCAAGAGTGTGATGAGCCTAC 1860
Qy 1983 AGACAAGCCCTAGGAATCTGCTCAAGCTCAGCTGA 2018
Db 1861 AGACAAGCCCTAGGAATCTGCTCAAGCTCAGCTGA 1896
```

```
RESULT 5
AY399039
LOCUS
DEFINITION Homo sapiens HCM0091 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
VERSION AY399039
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1887)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
```

```
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2' (bases 1 to 1887)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
this sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source 1..1887
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1887
/locus_tag="HCM0091"
ORIGIN
```

```
Query Match 43.6%; Score 1264; DB 9; Length 1887;
Best Local Similarity 78.5%; Pred. No. 7.1e-287;
Matches 1461; Conservative 0; Mismatches 380; Indels 21; Gaps 2;
Qy 169 CTTCCGCGCTCGGCTCGGAGAGCGCCGCGCTGGCGGAGGCTCATCAAGTCGCCTA 228
Db 35 CCACTCCCTCGCGCGAGAGCTAGCCCGCGCTGGCGGAGGCTCATCAAGTCGCCTA 94
Qy 229 AACCTCTGATGAAGAGCAGCGGTGAAGCGGACCATCAAAACACAACTCTGGGCACC 288
Db 95 AGCCCTATGAAGAGCAGCGGTGAAGCGGACCATCAAAACACAACTCTGGGCACC 154
Qy 289 GCTACAGATTCCTGGAGACGCTGGGCAAGGCGCCTTACGGAAGGTGAAGAGCGAG 348
Db 155 GCTACAGATTCCTGGAGACGCTGGGCAAGGCGCCTTACGGAAGGTGAAGAGCGAG 214
Qy 349 AGAGCTCGGGGCTCTGGTGGCCATCAAGTCCATCAGGAAGCAAAATCAAGATGAGC 408
Db 215 AGAGCTCGGGGCTCTGGTGGCCATCAAGTCCATCAGGAAGCAAAATCAAGATGAGC 274
Qy 409 AGGATCTGTGCAATACGAGGAGGAGATTGAGATCATGTCTTCACTCAACCCACCA 468
Db 275 AGGATCTGTGCAATACGAGGAGGAGATTGAGATCATGTCTTCACTCAACCCACCA 334
Qy 469 TCAATGCCATCCATGAAGTGTGTTGAGATAGCAGCAAGATTGTGATGATGAGATG 528
Db 335 TCAATGCCATCCATGAAGTGTGTTGAGATAGCAGCAAGATTGTGATGATGAGATG 394
Qy 529 CCAGCGAGGCGATCTGTATGATTACATCAGTGGGCGCCACGGCTGAGTGGGAGC 588
Db 395 CCAGCGAGGCGATCTGTATGATTACATCAGTGGGCGCCACGGCTGAGTGGGAGC 454
Qy 589 CCAGCATTTCTTCCGACAGATCTGTCTGCGCTGCACTACTGCCACAGAAAGGATCG 648
Db 455 CTAGGCAATTTCTTCCGCGAGATCTGTCTGCGCTGCACTACTGCCACAGAAAGATG 514
Qy 649 TTCAACCGAGATCTCAAGCTGGAAAAATCTTCTTCTAGATGCCAATGGAAACATCAAGATTG 708
Db 515 TCCACCGAGATCTCAAGCTGGAAAAATCTTCTTCTTGGATGCCAATGGAAATATCAAGNNNN 574
Qy 709 CTGACTTTGGCCTCTCCAACCTGTACCAAAAGGCAAGTTCCTCCAGACGCTTCTGTGGGA 768
Db 575 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 634
Qy 769 GCCTCTCTACGCTCGCTCGATAGTAGTCAACGGGAAGCCCTATGTGGGCGCCAGAGTGG 828
Db 635 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 694
Qy 829 ACAGCTGGTCTCTGGCGGTTCTCTCTGTACATCTCTGTGATGCCACCATGCCCCCTTTGACG 888
Db 695 ACAGCTGGTCTCTGGCGGTTCTCTCTGTACATCTCTGTGATGCCACCATGCCCCCTTTGATG 754
```



```
QY 889 GGCAGGATCATAAACACTGCTGAAGCAAAATCAGTAACAGGGCTTACCGTGAGCGGCCA 948
Db 755 GGCATGACCATAGATCTTAGTGAAACAGATCAGCAACGGGGCTACCGGAGCCACTTA 814
QY 949 AGCCGTCGAGTGCCTGTGGCTGTATCCGGTGGCTTTAAATGGTGAACCCCAACCGTCGGG 1008
Db 815 AACCTCTGATGCTGTGGCTGTATCCGGTGGCTTTGATGGTGAACCCCAACCGCGGG 874
QY 1009 CCACACTGGAGATGTAGCAGTCAATTTGGTGGTCAACTGGGGTTACACACCGGAGTCG 1068
Db 875 CCACCTTGAGGATGTGGCAGTCACTGTGGTCAACTGGGGTACGACCCACCGAGTGG 934
QY 1069 GGGAAACAGGAGCCCTCGCTGAGGCTGGCAACCTAGTGTGACTTTGGCGGGCCCTCCA 1128
Db 935 GAGGACAGGAGGCTCCCATGAGGCTGGCAACCTGGCAGTGAATCTGCCCGGCCCTCCA 994
QY 1129 TGGCGGACTGTTACGTGCTCTCCGCGCCCTCTCTGGAGAAATGGAGCAACGAGTGTGCA 1188
Db 995 TGGCTGACTGGCTCGGCGTTCTCTCCGCGCCCTCTCTGGAGAAATGGGCGCAAGTGTGCA 1054
QY 1189 GCTTCTTCAAGCAGCAGTGGCGGAGGTGAAGCACTGTACTGGGCTGGAGCGGCAAC 1248
Db 1055 GCTTCTTCAAGCAGCAGTGCACCTGTGGGGAAGCACCACCCCTGGGCTGGAGCGGCGAGC 1114
QY 1249 ATTCTCTTAAGAGTCCCGAAAGGAGAAATGACATGGCTCAAAATCTCAAGGTGACCCGG 1308
Db 1115 ATTCTCTTAAGAGTCCCGAAGGAGAAATGACATGGGCGGCTCTCTCCAGTGAACAGG 1174
QY 1309 CTGAGGATACCTTTCTCGCCCTGGCAAGAGCAGCCTTAAAGCTTCGAAAGGCAATTCOA 1368
Db 1175 CTGATGACACTGGCCATCGCCCTGGCAAGAGCAACCTCAAGTGCACCAAGGCAATTCOA 1234
QY 1369 AGAAAAGTCTTACCTCTGTCAGGAGGAGTACAGGAGGACCTCAGAACTCAGACCGG 1428
Db 1235 AGAAGAGGTTGACGCTCTGCAAGAGGTTACAGGAGGACCTTCGAGCTCAGGCCAA 1294
QY 1429 TGCTGTATCTCCAGGCGAGCTCTCTCTGCTGTATCTCTCTCCCAAGGAAGGCAATCC 1488
Db 1295 TCCCTGGAGCCGAGGAGGCTGCCCC-----CCCTGCTCTCCCAAGGAGGCAATTC 1345
QY 1489 TTAAGAGTCTGACAGCGTGAATCTGGTTACTACTCTCTCCAGAGCCCGAGGCTG 1548
Db 1346 TCAAGAGCCCGACAGCGAGTCTGGCTACTACTCTCTCTCCGAGCCCAAGTGAATCTG 1405
QY 1549 GGGNACTTTAGACGCCAGTGTGTGTGAGTGGGACCCCGTGGGAGGAGTCTC 1608
Db 1406 GGGAGCTCTTGGACGACAGGCGACGTGTGTGTGAGTGGGATCCCAAGGAGCAGAAGCCTC 1465
QY 1609 CACAGGCTTCAGGCTCTCTCTCCAGCGCAGGCGATTTCTCAAACTCAATGGCAAGTCT 1668
Db 1466 GCGNAGCTTCAGGCTGCTCTCTCATGCAAGGCAATCTCAAACTCAATGGCAAGTCT 1525
QY 1669 CCGCACAGCTTTAAGAGGCACTACCCCTAGACACTTTTGGCTCCCTGGACCAACTGGCCT 1728
Db 1526 CCCAGACAGCTTGGAGCTCGCGGCCCCCAACCACTTTCGGCTCCCTGGATGAATCGGCC 1585
QY 1729 CTTCCCATCTGACGCGCGGCCAGCGCCCTCAGGGGCTGTGAGTGAAGCAGCATCC 1788
Db 1586 CACCTCGCCCTCTGGCCCGGCGGACGACCTCAGGGGCTGTGAGGAGGAGCAGCATCC 1645
QY 1789 TGTCTCCGAGTCTTTGACCAATTTGACTTGGCTGAACGCTTCTCCGGAACCCCACTGA 1848
Db 1646 TGTCTCTGAGTCTTTGACCAAGTGAATTTGCTGAAAGGCTCTCCAGAGCCCACTGC 1705
QY 1849 GGGCTGTGTGTGTGTGACAACTGAGGGGGCTTGAGAGCCTCTCTCTCAAGAG----- 1902
Db 1706 GGGCTGTGTGTGTGTGACAACTGACGGGGCTTGAGAGCCTCTCTCTCAAGAGCCCTCTCAGAGGGCCCTG 1765
QY 1903 -----GTCCTGAAGGATGGTGGCAGGAACTCTTTGGGGGATGAGTCTTTCTCTGACAG 1956
Db 1766 GAAGCTGCCTGAGGCGCTGGCGGAGGATCCTTTGGGGGAGCAGTGTCTTTTCCCTGACAG 1825
```

```
QY 1957 ACTGCCAAGAGGTGACTGTCAGAGCCTCAGACAAAGCCCTAGGAATCTGCTCAAAGCTCAGCT 2016
Db 1826 ACTGCCAAGAGGTGACAGCGACCTCCGACAGGCACTGAGGGTCTGCTCAAAGCTCAGCT 1885
QY 2017 GA 2018
Db 1886 GA 1887

RESULT 6
AY399040 1887 bp DNA linear GSS 12-DEC-2003
LOCUS Pan troglodytes HOMO091 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY399040
VERSION AY399040.1 GI:39755029
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
MAMMALIA; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
REFERENCE 1 (bases 1 to 1887)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrieres,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1887)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrieres,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
source 1..1887
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
gene <1..>1887
/locus_tag="HCM0091"
ORIGIN
Query Match 32.8%; Score 951.4; DB 9; Length 1887;
Best Local Similarity 60.1%; Pred. No. 4e-213;
Matches 1119; Conservative 0; Mismatches 722; Indels 21; Gaps 2;
```

```
Qy 469 TCATTGTCATCCATGAATGTTTGAAGATAGACAGAAATTGTGTTGTCATGAGTATG 528
Db |||||
Qy 335 TCATTGTCATCCATGAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 394
Db
Qy 529 CCAGCCGAGGATCTGTATGATTACATCAGTGAAGCCGACGGCTGAGTGAGCGGACG 588
Db
Qy 395 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 454
Qy 589 CCAGGCATTTCTCCGACAGATCGTGTGCTGCCCTGCACTACTGCCACCAAGCGGATCG 648
Db |||||
Qy 455 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 514
Qy 649 TTACCCGAGATCTCAAGCTGGAAGAAACATCTCTTAGATGCAATGGAACATCAAGATTG 708
Db |||||
Qy 515 TCCACCGAGATCTCAAGCTGGAAGAAACATCTCTTTGGATGCCAATATCAAGNNNN 574
Qy 709 CTGACTTTGGCCTCTCAACTGTATCAACAAAGCAAGTTCTCCAGACGTTCTGTGGGA 768
Db
Qy 575 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 634
Qy 769 GCCCTCTCTACGCCCTCGCTGAGATAGTCAACGGAAGCCCTATGTGGGCCAGAGTGG 828
Db
Qy 635 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 694
Qy 829 ACAGCTGCTCTCTGGCGCTTCTCTGTATCATCTCTGTGTCATGGCACCATGCCCTTTGACG 888
Db
Qy 695 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 754
Qy 889 GGCAGGATCAATAAAGCTGTGAAGCAATCAGTAAGGGGCTTACCGTGAGCGGCCCA 948
Db
Qy 755 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 814
Qy 949 AGCCGTCGATGCTGGCCTGATCGGCTGCTGATGCTGTAAGTGAACCCACCGCTCGGG 1008
Db
Qy 815 NNNNNNNNNATGCTGTGGCCNGATCCGGTGGCTGTTGATGGTGAACCCACCGCCGG 874
Qy 1009 CCACACTGGAGATGTAGCCAGTCAITGGTGGGTCAACTGGGGTTACACCACCGGAGTCG 1068
Db |||||
Qy 875 CCACCTTGGAGGATGTGGCCAGTCACTGTTGGTCAACTGGGCTAGCCACCCGAGTGG 934
Qy 1069 GGGAAAGGAGGCTCGCTGAGGGTGGGACCCCTAGTGTGACTTTGGCCGGGCTTCCA 1128
Db |||||
Qy 935 GAGAGCAGGAGGCTCCGATGAGGGTGGGACCCCTGGCAGTACTCTGTCCGGGCTCCG 994
Qy 1129 TGGGGACTGCTTACGCTCTCTCGGCCCTCTCTGGAGGATGGAGCAAGGTGCA 1188
Db |||||
Qy 995 TGGCTGACTGGCTCGGCTTCTCCGCCCTCTCTGGAGGATGGGCAAGGTGCA 1054
Qy 1189 GCTTCTTCAAGCAGCAGTCCGGGAGGTGGAAGCACTGTACCTGGGCTGGAGCGGCAAC 1248
Db |||||
Qy 1055 GCTTCTTCAAGCAGCAGTCCGCTGGTGGGGAAGCACCCTGGCTGGAGCGGAGC 1114
Qy 1249 ATTCTTTAAGAGTCCGGAAGAGAGATGACATGGCTCAAAATCTGCAAGGTGACCCGG 1308
Db |||||
Qy 1115 ATTCTGCTCAAGAGTCCGGAAGAGAGATGACATGGCCCACTCTCTCACAGTGACACGG 1174
Qy 1309 CTGAGGATACCTCTCTCGCCCTGGCAAGAGAGCTTAAAGTCTCGAAGGATTTCTCA 1368
Db |||||
Qy 1175 CTGATGACTGCCCCCTGGCAAGAGAGACTTCAAGTGTGCCAAGGGGATTTCTCA 1234
Qy 1369 AGAAAAGTCTCTTACCTCTGAGGGAGGTACAGGAGGACCTCTCAGAACTCAGACCGG 1428
Db |||||
Qy 1235 AGAAGAGGTGTGAGCTCTGAGAGGGGGGACAGAGGACCTCTCGAGCTCAGCCCA 1294
Qy 1429 TGCTGTATCTCAGGGCAGCTGTCTCTGCTGTATCTCTGCTCCCAAGAAAGGATCTC 1488
Db |||||
Qy 1295 TCCCTGCGAGCCAGGCGAGCTGCC-----CCCTGCTCTCCCAAGAGGCGATTC 1345
Qy 1489 TTAAGAGTCTGACAGCGGTGAATCTGTTACTACTCTCTCAGAGGCCAGCGAGTCTG 1548
Db |||||
Qy 1346 TCAAGAGCCCAACAGCGGAGTGTGGCTACTACTCTCTCCGAGGCCAGTGAATCTG 1405
Qy 1549 GGGAACTCTTAGAGCCGAGTGTGTTGTGAGTGGGGACCCCGCTGGAGCAGAGTCTC 1608
```

```
Db |||||
Qy 1406 GGAGCTCANNAGACGAGCGGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1465
Qy 1609 CACAGGCTTCAGGCTCTCTCTCCACGCGCAGGGCATTTCTCAAACTCAATGGCAAGTCT 1668
Db
Qy 1466 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1525
Qy 1669 CCGCAGCAGCTTAGAAGGCACTACCCCTAGCACCTTTGGCTCCCTGGACCAACTGGCCT 1728
Db |||||
Qy 1526 NNNNGACAGCTTGGAGCTCGCGGCCCCACACCTTTGGCTCTTGGATGAATCGCCC 1585
Qy 1729 CTCCCATCTGTGAGCCCGGCCAGCCCTCAGGGGCTGTGAGTGAGGACAGCATCC 1788
Db |||||
Qy 1586 CATCTCGCCCTGGCGCGGCCAGCCCTCAGGGCTGTGAGGAGGACAGCATCC 1645
Qy 1789 TGTCTCTCCGAGTCTTTGACCAATTTGACTTGGCTGAACTCTTCCGAAACCCCACTGA 1848
Db |||||
Qy 1646 TGTCTCTGAGTCTTTGACCAATTTGACTTGGCTGAACTCTTCCGAAACCCCACTGC 1705
Qy 1849 GGGCTGTGTCTGTGGCAACCTGAGGGGCTTGGAGCAGCTCTCTCAGAAG----- 1902
Db |||||
Qy 1706 GGGCTGTGTCTGTGGCAACCTTATGGGCTTGGAGGAGCCCTCAGAGGGCCCTG 1765
Qy 1903 -----GTCTGAAGCGATGTGGCAGGAATCTTGGGGGATAGCTCTTTCTTGACAG 1956
Db |||||
Qy 1766 GAAGCTGCTGAGCGCTGGCGGAGATCTTTTGGGGGACAGCTGCTTTTCTTGACAG 1825
Qy 1957 ACTGCCAAGAGTGTGCTGAGCCTCAGACCAAGCCCTAGGAATCTGCTCAAGCTCAGCT 2016
Db |||||
Qy 1826 ACTGCCAGGAGGTGACAGCAGCTTACCGGAGGCACTGAGGGTCTGCTCAAGCTCACCT 1885
Qy 2017 GA 2018
Db ||
Qy 1886 GA 1887
```

## RESULT 7

BI653092

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI653092 746 bp mRNA linear EST 12-SEP-2001  
603300616F1 NCI\_CGAP\_Mam3 Mus musculus cDNA clone IMAGE:5341299 5',  
mRNA sequence.  
BI653092  
BI653092.1 GI:15567328  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NIH-MGC <http://mgi.nhl.nih.gov/>  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLML1866 row: p column: 04  
High quality sequence stop: 746.  
Location/Qualifiers  
1. .746  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="129, C57BL/6J.FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5341299"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="10 months"  
/lab\_host="DH10B"

## FEATURES

source

1. .746

/clone.lib="NCI CGAP Mam3"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigators  
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH  
 Reference for transgenic model: Xu et al., Nature Genetics  
 22, 37-43 (1999)."

## ORIGIN

Query Match 24.8%; Score 721; DB 4; Length 746;  
 Best Local Similarity 98.0%; Pred. No. 7.7e-159;  
 Matches 730; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1787 CCTGCTCTCCAGTCCCTTTGACCAATGGACTTCCTGGAAGCTTCCCGAAGCCCACT 1846  
 Db 2 CCACGCTGCCCCACGCGTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 61  
 QY 1847 GAGGGCTGT 1906  
 Db 62 GAGGGCTGT 121  
 QY 1907 GAACGATGTGGCAGGAATCCTTGGGGATAGTCTGCTTTCTCTGACAGACTGCCAAGA 1966  
 Db 122 GAACGATGTGGCAGGAATCCTTGGGGATAGTCTGCTTTCTCTGACAGACTGCCAAGA 181  
 QY 1967 GGTGACTGCGACCTACAGAACCCCTAGGAATCTGCTCAAAGCTCAGCTGAGGAAGGA 2026  
 Db 182 GGTGACTGCGACCTACAGAACCCCTAGGAATCTGCTCAAAGCTCAGCTGAGGAAGGA 241  
 QY 2027 GATGGTCCCTAGTATGGGTAGGCTCTGAGAGGGTTTTCAGAGAGAACCCCTGGTCCGAT 2086  
 Db 242 GATGGTCCCTAGTATGGGTAGGCTCTGAGAGGGTTTTCAGAGAGAACCCCTGGTCCGAT 301  
 QY 2087 TCCTCCAGTGAATAGATACATCAAGGCTCTAGTCTGAGGCTGAGTCACTGAGTGAAG 2146  
 Db 302 TCCTCCAGTGAATAGATACATCAAGGCTCTAGTCTGAGGCTGAGTCACTGAGTGAAG 361  
 QY 2147 ATGAGAGAAATCGCATTTGATGTGAAAGGAATGGAACCCCTTGTGCTGCCGAGTGTATAG 2206  
 Db 362 ATGAGAGAAATCGCATTTGATGTGAAAGGAATGGAACCCCTTGTGCTGCCGAGTGTATAG 421  
 QY 2207 TGGGTGGCTGGAAGTGCCTACCTCTTTGTGCCATGAGTGCACCATGACATTTCCC 2266  
 Db 422 TGGGTGGCTGGAAGTGCCTACCTCTTTGTGCCATGAGTGCACCATGACATTTCCC 481  
 QY 2267 ACCGTGTTCTGCTGCTGACCTTACATAGATTTCTGTTTCCATCAACACCGAGTTAG 2326  
 Db 482 ACCGTGTTCTGCTGCTGACCTTACATAGATTTCTGTTTCCATCAACACCGAGTTAG 541  
 QY 2327 AACCTGACTTTCCTGGGAGGTATGTGTAGTACTGCCATTATTTAGAGAGGAACAGCC 2386  
 Db 542 AACCTGACTTTCCTGGGAGGTATGTGTAGTACTGCCATTATTTAGAGAGGAACAGCC 601  
 QY 2387 TCTGGTTTCCATCTCTGCTGTGCACTCAAGACCTGGGAAGACTCGGACCGTGT 2446  
 Db 602 TCTGGTTTCCATCTCTGCTGTGCACTCAAGACCTGGGAAGACTCGGACCGTGT 661  
 QY 2447 TGACTTCACTCAAGGGGACAGATGCCCTGGACCCCATCTTAGATCTCAGAGACTTGA 2506  
 Db 662 TGACTTCACTCAAGGGGACAGATGCCCTGGACCCCATCTTAGATCTCAGAGACTTGA 721  
 QY 2507 ACCTTGAAGCTGTTTCTTAGTACCCA 2531  
 Db 722 ACCTTGAAGCTGTTTCTTAGTACCCA 746

## RESULT 8

CA319312  
 LOCUS  
 DEFINITION  
 UI-M-FW0-cbz-i-16-0-UI.r1 NIH BMAP\_FW0 Mus musculus cdna clone  
 CA319312 795 bp mRNA linear EST 09-JUL-2003  
 IMAGE:6816569 5', mRNA sequence.  
 CA319312  
 ACCESSION  
 CA319312.1 GI:24537436

## KEYWORDS

SOURCE  
 ORGANISM

## EST.

Mus musculus (house mouse)  
 Mus musculus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 795)

NIH-MGC http://mgi.nci.nih.gov/.

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. .795

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6"

/db\_xref="taxon:10090"

/clone="IMAGE:6816569"

/tissue\_type="whole brain"

/dev\_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"

/lab\_host="DH10B (T1 phage resistant)"

/clone\_lib="NIH BMAP\_FW0"

/notes="Organ: Brain; Vector: pYX-Aec; Site 1: EcoR I;

Site 2: Not I; The library was constructed according

Bonaldi, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Aec vector. The library tag

is AGCGAGACAG. This library was created for the polyA tail

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

program coordinator."

## ORIGIN

Query Match 24.8%; Score 720; DB 6; Length 795;  
 Best Local Similarity 96.8%; Pred. No. 1.3e-158;  
 Matches 765; Conservative 0; Mismatches 0; Indels 25; Gaps 2;

QY 43 GCGGCTGTGACCTCTGAGCCCGCGCTCAGCGCGCTGCTACTGTGCTGCCGACCCACTC 102

Db 1 GCGGCTGTGACCTCTGAGCCCGCGCTCAGCGCGCTGCTACTGTGCTGCCGACCCACTC 60

QY 103 CACCTCGCGTCCCGCACCATGGAGTGGTGGCTTACTCCAGCGCCCGGAGGCTC 162

Db 61 CACCTCGCGTCCCGCACCATGGAGTGGTGGCTTACTCCAGCGCCCGGAGGCTC 120

QY 163 CTTGCGCTCCGCTTGGCTCGGAGAGCGCCGCGCTGGCGGACGGGCTCATCAAGT 222

Db 121 CTTGCGCTCCGCTTGGCTCGGAGAGCGCCGCGCTGGCGGACGGGCTCATCAAGT 180

QY 223 CCGCTAAACCTCTGATGAAGAAGCGGCTGAAGCGGACCATCATCAAAACAAACCTGC 282

Db 181 CGCCTAAACCTCTGATGAAGAAGCGGCTGAAGCGGACCATCATCAAAACAAACCTGC 240

QY 283 GGCACCGCTACAGTTCCTGGAGAGCGTGGCAAGGACCTACGGGAGGTGGAAGG 342

Db 241 GGCACCGCTACAGTTCCTGGAGAGCGTGGCAAGGACCTACGGGAGGTGGAAGG 300

```
QY 343 CACGAGAGAGCTCGGGCGCTGTGTGGCCATCAAGTCCATCAGGAAGACAAAATCAAAG 402
Db 301 CACGAGAGAGCTCGGGCGCTGTGTGGCCATCAAGTCCATCAGGAAGACAAAATCAAAG 360
QY 403 ATGAGCAGGATCTGCTGCATACATACGAGGAGATGAGATCATGTCTTCACTCAACACC 462
Db 361 ATGAGCAGGATCTGCTGCATACATACGAGGAGATGAGATCATGTCTTCACTCAACACC 420
QY 463 CCACATCATTTGGCATCCATGAA-----GTGTTTGAGAAATA 498
Db 421 CCACATCATTTGGCATCCATGAAAGTGGGACATCACGCTAGTGACTGTGTTGAGAAATA 480
QY 499 GCAGCAAGATGTGATGTTCATGAGATATGCCAGCGAGCGGATCTGTATGATATCATCA 558
Db 481 GCAGCAAGATGTGATGTTCATGAGATATGCCAGCGAGCGGATCTGTATGATATCATCA 540
QY 559 GTGAGCGGCCACGCTGAGTGAGGCGGACGCCAGGCAATTTCTTCCGACAGATCGTGTCTG 618
Db 541 GTGAGCGGCCACGCTGAGTGAGGCGGACGCCAGGCAATTTCTTCCGACAGATCGTGTCTG 600
QY 619 CCCTGCACACTACTGCCACCAAGCGGATCGTTCAACGAGATCTCAAGCTCGAAACATCC 678
Db 601 CCCTGCACACTACTGCCACCAAGCGGATCGTTCAACGAGATCTCAAGCTCGAAACATCC 660
QY 679 TTCTAGATGCAATGGAACATCAAGATTCCTGATTTGGCCCTCTCAACCTGTACACACA 738
Db 661 TTCTAGATGCAATGGAACATCAAGATTCCTGATTTGGCCCTCTCAACCTGTACACACA 720
QY 739 AAGCAAGATTCCTGCAGACCTTCTGTGGAGCCCTCTACGCTCGCCCTGAGATAGTCA 798
Db 721 AAGCAAGATTCCTGCAGACCTTCTGTGGAG-CCTCTCTACGCTCGCCCTGAGATAGTCA 779
QY 799 ACCGGAAGCC 808
Db 780 ACCGGAAGCC 789
```

```
RESULT 9
CF729223
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CF729223 727 bp mRNA linear EST 09-OCT-2003
UI-M-HD0-clc-b-24-0-UI.r1 NIH_BMAP_HD0 Mus musculus cDNA clone
IMAGE:30615119 5', mRNA sequence.
CF729223
CF729223.1 GI:37603391
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 727)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
(This clone was contributed by the Brain Molecular Anatomy Project
(BMAP))
Seq primer: pYX-5.
Location/Qualifiers
1. .727
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30615119"
/tissue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
```

```
/lab host="DH10B (T1 phage resistant)"
/clone lib="NIH_BMAP_HD0"
/notes="Organ: Eye; Vector: pYX-Asc; Site: 1; Ecor 1;
Site 2: Not 1; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TTTTGAAGT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."
```

ORIGIN

Query Match	Best Local Similarity	24.6%	Score 714;	DB 7;	Length 727;				
Matches	725;	Conservative	0;	Mismatches	2;	Indels	1;	Gaps	1;
QY	1828	GTCTTCCGAAACCCACACCTAGAGGGGCTGTGTCTGTGGACACCTAGAGGGGCTTGAGC	1887						
Db	1	GTCTTCCGAAACCCACACCTAGAGGGGCTGTGTCTGTGGAC-ACCTAGAGGGGCTTGAGC	59						
QY	1888	AGCTCCCTCAGAGGCTCTGAAGCGATGGTGGCAGGAATCCTTGGGGGATAGCTGCTTTT	1947						
Db	60	AGCTCCCTCAGAGGCTCTGAAGCGATGGTGGCAGGAATCCTTGGGGGATAGCTGCTTTT	119						
QY	1948	CTCTGACAGATGTCGAAGAGGCTGACCTGCAGCCTACAGAACGCCCTAGGAATCTGCTCAA	2007						
Db	120	CTCTGACAGATGTCGAAGAGGCTGACCTGCAGCCTACAGAACGCCCTAGGAATCTGCTCAA	179						
QY	2008	AGCTAGCTGAGGAAGGGAGATGGTCCCTAGTATGGGGTAGGCTCTGAGAGGGTTTGCA	2067						
Db	180	AGCTAGCTGAGGAAGGGAGATGGTCCCTAGTATGGGGTAGGCTCTGAGAGGGTTTGCA	239						
QY	2068	GAGGAACCTTGGGTCGGATTTCCTCCAGTGAATAGATACATCAAGGCTCTACGCTCGCA	2127						
Db	240	GAGGAACCTTGGGTCGGATTTCCTCCAGTGAATAGATACATCAAGGCTCTACGCTCGCA	299						
QY	2128	GCCTGACTGAACCTGAAAGATGAGAAATCGCATTTGATGTGGAAGGAATGGGAACCT	2187						
Db	300	GCCTGACTGAACCTGAAAGATGAGAAATCGCATTTGATGTGGAAGGAATGGGAACCT	359						
QY	2188	TGCTGCCCGAGTGTATAGTGGGGTGGCTGAAGGTCCTACCTCTTTGTGCCATGAT	2247						
Db	360	TGCTGCCCGAGTGTATAGTGGGGTGGCTGAAGGTCCTACCTCTTTGTGCCATGAT	419						
QY	2248	GTCAACCATGACATTTCCCAACCTGTTCTGTGGCTGCACCTTCACATAAGTTTCTGTTTC	2307						
Db	420	GTCAACCATGACATTTCCCAACCTGTTCTGTGGCTGCACCTTCACATAAGTTTCTGTTTC	479						
QY	2308	CATCAACCAACCGGGTTAGAACCTTGACTTCTCTGGGAGGTAATGTGTAGTACTGCCATT	2367						
Db	480	CATCAACCAACCGGGTTAGAACCTTGACTTCTCTGGGAGGTAATGTGTAGTACTGCCATT	539						
QY	2368	ATTAGAGAGGAAACAGCTCTGGTTTCATCTCTGCTGTGTGATCTCAAGAACCTGG	2427						
Db	540	ATTAGAGAGGAAACAGCTCTGGTTTCATCTCTGCTGTGTGATCTCAAGAACCTGG	599						
QY	2428	GAAAGACTCGGACCGCTGTTTGAATCTCTCAAGGGGACAGATGCCCTGGACCCCATC	2487						
Db	600	GAAAGACTCGGACCGCTGTTTGAATCTCTCAAGGGGACAGATGCCCTGGACCCCATC	659						
QY	2488	TTAGATCTCAGAGACTTGAACCTTGAAGCTGTTTCTTAGTACCAGATGTGGATGGTCT	2547						
Db	660	TTAGATCTCAGAGACTTGAACCTTGAAGCTGTTTCTTAGTACCAGATGTGGATGGTCT	719						
QY	2548	CTGTTTCT 2555							
Db	720	CTGTTTCT 727							

```

RESULT 10
BG915967
LOCUS
DEFINITION
  BG915967 602815231F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4937647 5',
  mRNA sequence.
ACCESSION
  BG915967
VERSION
  BG915967.1 GI:14296443
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  NIH-MGC http://mgi.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
  Ph.D.
  CDNA Library Preparation: Life Technologies, Inc.
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LHAM10871 row: m column: 08
  High quality sequence stop: 695.
  Location/Qualifiers
    1..992
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="NMRI"
      /db_xref="taxon:10090"
      /clone="IMAGE:4937647"
      /tissue type="tumor, gross tissue"
      /dev stage="5 months"
      /lab_host="DH10B"
      /clone_lib="NCI_CGAP_Mam4"
      /note="Organ: mammary; Vector: pCMV-SPORT6; Site1: SalI;
      Site2: NotI; Cloned unidirectionally. Primer: Oligo dr.
      Library constructed by Life Technologies. Investigators
      providing samples: Lothar Hennighausen/Priscilla Furth,
      NIH Reference for transgenic model: Li et al., Cell Growth
      and Differentiation 7, 3-11 (1996)."
```

## ORIGIN

```

Query Match      24.1%; Score 700.2; DB 4; Length 992;
Best Local Similarity 92.2%; Pred. No. 6.6e-154;
Matches 857; Conservative 0; Mismatches 60; Indels 12; Gaps 11;

QY 1931 GGGGATAGCTGCTTTCTCTGACAGACTGCCAAGAGTGACTGCAGCCTACAGCAAGC 1990
DB 13 GGGGATAGCTGCTTTCTCTGACAGACTGCCAAGAGTGACTGCAGCCTACAGCAAGC 72

QY 1991 CCTAGGAATCTGCTCAAGCTCAGCTCAGGAAGGAGATGCTGCCCTAGTATGGGTAGG 2050
DB 73 CCTAGGAATCTGCTCAAGCTCAGCTCAGGAAGGAGATGCTGCCCTAGTATGGGTAGG 132

QY 2051 CTCTGAGAGGTTTGCAGAGGAACCCCTGGTCCGATTCCTCCAGTGAATAGATACATCA 2110
DB 133 CTCTGAGAGGTTTGCAGAGGAACCCCTGGTCCGATTCCTCCAGTGAATAGATACATCA 192

QY 2111 AGGGCTCTAGCTGACGCTGACTGACCTGAAAGATGAGAAATCGAATGATGTGG 2170
DB 193 AGGGCTCTAGCTGACGCTGACTGACCTGAAAGATGAGAAATCGAATGATGTGG 252

QY 2171 AAAGGAATGGAAACCCCTGCTGCCCGAGTGTATAGTGGGGTGGCTGAAGGTGCCTACC 2230
DB 253 AAAGGAATGGAAACCCCTGCTGCCCGAGTGTATAGTGGGGTGGCTGAAGGTGCCTACC 312
```

```

QY 2231 TCCTTTGTGCGCATGAGTGTACCCATGACATTTTCCACCCCTGTCTCTGCTGCACCTTC 2290
DB 313 TCCTTTGTGCGCATGAGTGTACCCATGACATTTTCCACCCCTGTCTCTGCTGCACCTTC 372

QY 2291 ACATAAGTTTCTGTTTCCATCAACCAAGGGTTAGAACCTGACATTCCTGGGAGGTA-A 2349
DB 373 ACATAAGTTTCTGTTTCCATCANACA-CAGGGTTAGAACCTGACATTCCTGGGAGGTACA 431

QY 2350 TGTGTAGTGAATGCCATTTATTTAGAGAGGAAACAGACCTCTGGTTTCCATCTCTGCTGCTG 2409
DB 432 TGTGTAGTGAATGCCATTTATTTAGAGAGGAAACAGACCTCTGGTTTCCATCTCTGCTGCTG 491

QY 2410 TGCAATCTCAAGACCTGGGAGAGCTCGGACCGCTGTTTGAATTCATCTCAAGGGGACCAG 2469
DB 492 TGCAATCTCAAGACCTGGGAGAGCTCGGACCGCTGTTTGAATTCATCTCAAGGGGACCAG 551

QY 2470 ATGCCCTCTGACCCCATCTTAGATCTCAGAGACTTGAACCTTGAAGCTGTTCTAGTAGTACC 2529
DB 552 ATGCCCTCTGGA-CCCATCTTAGATCTCAGAGACTTGAACCTTGAAGCTGTTCTAGTAGTACC 610

QY 2530 CAGATGTGGA-TGGATGCTCTGTTTCTCAGGCCAACCGGACCTAGAAATGTGCTGACTTAT 2588
DB 611 CAGATGTGGAATGATGCTCTGTTTCTCAGGCCAACCGGACCTAGAAATGTGCTGACTTAT 670

QY 2589 TTATTTTGTGATTCCTCACTTCTGTTTGTGTTT--TTTGTGTTTGTGTTTGTGTTT 2646
DB 671 NTAATCTCCCGTGATTCCTCACTTCTGTTTTCAGGTCAAAGATCGAGCCGTAGCCCAAGG 730

QY 2647 TTTTAAAGTGAATTTTGTCT-GCTTTCAATTAATGTGAATGCTGTGTTCTGGGAACTCCAC 2705
DB 731 TCCTAAGTGCATCTTGTCTGGCTTTCACCTACTGTAATGCTGTGTTCTGGGG-ACTCCAC 789

QY 2706 TGTGCCACTGAAGTTTATGTACAGAAATGATTTGGCAATGATG-TCCCTCTATTTCAGG 2764
DB 790 TGTGCC-CTGAAGTCTAGGTCCGGCGAAAGTTTGGCACTGATGTTCCCTCTATTTCAGG 848

QY 2765 GGGGTGGGGCGGTTTTTCA-AATGATGCTCTGAGCAGCTGCTGGATTGAGTCTCCAGTC 2823
DB 849 GGGGTGGGGCGGTTTTTCAATGCTGCTGTTTGAAGCAGCTGGTGGATTGCTT-TCCTGTC 907

QY 2824 CTTTACACCCCAAGGCTGCCACCCCTCC 2852
DB 908 CTTTCAGGCCCAAGGTGGGCCACCCCTGCC 936

RESULT 11
CF726196
LOCUS
DEFINITION
  UI-M-GZO-cjhp-h-08-0-UI.r1 NIH BMAP_GZO Mus musculus cDNA clone
  IMAGE:30606415 5', mRNA sequence.
ACCESSION
  CF726196
VERSION
  CF726196.1 GI:37600364
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  NIH-MGC http://mgi.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: Dr. James Lin University of Iowa
  CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
  CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Distribution information can be found at
  http://genome.uiowa.edu/distribution/mousefl.html
  This clone was contributed by the Brain Molecular Anatomy Project
  (BMAP)
  Seq primer: pYX-5.
  Location/Qualifiers
```

## FEATURES

```
source
1. 705
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30606415"
/tissue_type="whole eye"
/dev_stage="embryo 12.5, 13.5, 14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_G20"
/notes="Organ: Eye; Vector: pYX-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with Not I, and then directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TTATGTAAGT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
```

ORIGIN

Query Match 23.7%; Score 688; DB 7; Length 705;  
Best Local Similarity 99.3%; Pred. No. 4.6e-151;  
Matches 688; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 31 GCGTGTCTGGTGGCGGTGTGACCTCTGAGCCCGGCTCAGCGCGGCTGCTACTGCTG 90  
DB 13 GCGTGTCTGGTGGCGGTGTGACCTCTGAGCCCGGCTCAGCGCGGCTGCTACTGCTG 72

QY 91 CCCGACCACTCCACCTCGCGGTCCCGCCACCATGAGTGGTGGCTTACTCCAGCGCC 150  
DB 73 CCCGACCACTCCACCTCGCGGTCCCGCCACCATGAGTGGTGGCTTACTCCAGCGCC 132

QY 151 CGAGCCAGGCTCCCTCGGCTCCGCTCGGCTCGGCTCGGAGAGCGCCCGCGCTGCGGAGC 210  
DB 133 CGAGCCAGGCTCCCTCGGCTCCGCTCGGCTCGGAGAGCGCCCGCGCTGCGGAGC 192

QY 211 GGCTCATCAAGTCCCTAAACCTCTGATGAGAGCAGCGGCTGAAGCGGACCATCACA 270  
DB 193 GGCTCATCAAGTCCCTAAACCTCTGATGAGAGCAGCGGCTGAAGCGGACCATCACA 252

QY 271 AACACAACTCGCGCACCGCTTACGAGTTCTTGAGAGCGCTGGGCAAGGCACTTACGGGA 330  
DB 253 AACACAACTCGCGCACCGCTTACGAGTTCTTGAGAGCGCTGGGCAAGGCACTTACGGGA 312

QY 331 AGGTGAAGAGGACGAGAGAGCTCGGGCGTCTGTGGGCAATCAAGTCCATCAGGAAG 390  
DB 313 AGGTGAAGAGGACGAGAGAGCTCGGGCGTCTGTGGGCAATCAAGTCCATCAGGAAG 372

QY 391 ACAAATCAAGATGACGAGATCTGTGCATACGAGGAGGATGAGATCATGCTT 450  
DB 373 ACAAATCAAGATGACGAGATCTGTGCATACGAGGAGGATGAGATCATGCTT 432

QY 451 CACTCAACCAACCCCAATCATTCGCCATCATGAAGTGTGAGATGAGCAAGATG 510  
DB 433 CACTCAACCAACCCCAATCATTCGCCATCATGAAGTGTGAGATGAGCAAGATG 492

QY 511 TGATTGTATGAGATGATCCAGCGGATCTGTATGATTAATCATCAGTGAAGCGGAC 570  
DB 493 TGATTGTATGAGATGATCCAGCGGATCTGTATGATTAATCATCAGTGAAGCGGAC 552

QY 571 GGCTGAGTGAGCGGAGCGGAGGATTTCTTCCAGAGATGCTCTGCGCTGCACTACT 630  
DB 553 GGCTGAGTGAGCGGAGCGGAGGATTTCTTCCAGAGATGCTCTGCGCTGCACTACT 612

QY 631 GCCACCAAGAGGATCGTTCCAGCGATCTCAAGCTGGAAACATCTCTTAGATGCA 690

Db 613 GCCACCAAGAGGATCGTTCCAGCGATCTCAGCGATCTCANGCTGAANACATCTCTTAGATGCA 672

QY 691 ATGGAACATCAAGATGCTGACTTTGGCTCT 723

Db 673 NTGGAACATCAAGATGCTGACTTTGGCTCT 705

RESULT 12

LOCUS CB248251 726 bp mRNA linear EST 15-JUL-2003

DEFINITION UI-M-FD0-byi-a-21-0-UI.r1 NIH\_BMAP\_F00 Mus musculus cDNA clone IMAGE:5718428 5', mRNA sequence.

ACCESSION CB248251

VERSION CB248251.1 GI:28386432

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 726)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
Cloning Distribution: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. 726  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5718428"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 12.5 dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH\_BMAP\_F00"  
/notes="Organ: brain; Vector: pYX-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAGAGGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 23.6%; Score 683.6; DB 6; Length 726;  
Best Local Similarity 98.8%; Pred. No. 5.1e-150;  
Matches 720; Conservative 0; Mismatches 5; Indels 4; Gaps 3;

QY 922 GTAAACGGGGTTAC-CGTGAGCGCCCAAGCCCTCCGATCGCTGCGCTGATCCGGTGG 980  
DB 1 GTAAACGGGGTTACNCNGTGAAGCGCCCAAGCCCTCCGATCGCTGCGCTGATCCGGTGG 60

QY 981 CTGTTAATGTGAACCCCGCTCGGGGCCACACTGAGGAGTGTAGCCAGTCTATTGTTGG 1040



```

Db      61  CTGTTAATGTGAACCCACCCGTCGGGCCACACTGGAGGATGTAGCAGTCATTGGTGG 120
QY      1041 GTCAACTGGGTTACACACCGGAGTCGGGAAACAGAAAGCCCTCGTGGAGTGGGCAC 1100
Db      121  GTCAACTGGGTTACACACCGGAGTCGGGAAACAGAAAGCCCTCGTGGAGTGGGCAC 180
QY      1101 CCTAGTGGTGACTTTGGCGGGCTCCATGCGGACCTGGTTACGTGCTCTCGCGCCCC 1160
Db      181  CCTAGTGGTGACTTTGGCGGGCTCCATGCGGACCTGGTTACGTGCTCTCGCGCCCC 240
QY      1161 CTCCTGAGAAATGAGCAAGGTGTGAGCTTTCTCAAGCAGACAGTCGCCGGAGGTGGA 1220
Db      241  CTCCTGAGAAATGAGCAAGGTGTGAGCTTTCTCAAGCAGACAGTCGCCGGAGGTGGA 300
QY      1221 AGCACTGTACTGGCTGGAGCGGCAATCTCTTAAGAGTCCGAAAGGAGATGAC 1280
Db      301  AGCACTGTACTGGCTGGAGCGGCAATCTCTTAAGAGTCCGAAAGGAGATGAC 360
QY      1281 ATGGCTCAAAATCTGCAAGGTGACCGGCTGAGGATACCTCTTCTCGCCCTGGCAAGGC 1340
Db      361  ATGGCTCAAAATCTGCAAGGTGACCGGCTGAGGATACCTCTTCTCGCCCTGGCAAGGC 420
QY      1341 AGCCTTAAGTTCGAAAGGCAATCTCAAGAAAAGTCTCTACCTCGTCAGGGAGGTA 1400
Db      421  AGCCTTAAGTTCGAAAGGCAATCTCAAGAAAAGTCTCTACCTCGTCAGGGAGGTA 480
QY      1401 CAGGAGACCTCAGGACCTCAGACCGTGGCTGATCTACTCCAGGCGACCTGTCCCTGCT 1460
Db      481  CAGGAGACCTCAGGACCTCAGACCGTGGCTGATCTACTCCAGGCGACCTGTCCCTGCT 540
QY      1461 GTATCCCTGTCTCCCAAGGAAGGCAATCTTAAGAAAGTCTCAGACGCTGATCTGGTTAC 1520
Db      541  GTATCCCTGTCTCCCAAGGAAGGCAATCTTAAGAAAGTCTCAGACGCTGATCTGGTTAC 600
QY      1521 TACTCTCTCAGAGCCAGGAGTCTGGGAACTCTTAGACGCGCAGTGTGTTGTG 1580
Db      601  TACTCTCTCAGAGCCAGGAGTCTGGGAACTCTTAGACGCGCAGTGTGTTGTG 660
QY      1581 AGTGGGACCCGTTGAGGAGGAGTCTCCAGGCTTCAGGGCTCTCCCTCCACCGCAAG 1640
Db      661  AGCGNGA--CCGTGGAGCAGAA--TCTCAGGCTTCAGGGCTCTCTCCACCGCAAG 717
QY      1641 GGCATTCTC 1649
Db      718  GGCATTCTC 726

RESULT 13
LOCUS   CO424322
DEFINITION  UI-M-HUO-cqx-a-14-0-UI.r1 NIH_BMAP_HU0 Mus musculus cDNA clone
IMAGE:30665389 5', mRNA sequence.
ACCESSION  CO424322
VERSION    CO424322.1
KEYWORDS   GI:49670481
SOURCE     Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaops@email.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

```

```

FEATURES             source
Seq primer: pyX-5.   Location/Qualifiers
1..672
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30665389"
/tissue_type="whole eye"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_HU0"
/note="Organ: Eye; Vector: pyX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pyX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AATAATTACG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."
ORIGIN
Query Match      23.2%; Score 672; DB 7; Length 672;
Best Local Similarity 100.0%; Pred. No. 2,7e-147;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      913  AGCAATCAGTAACGGGGCTTACCGTAGCGCGCCCAAGCCGTCGATCGCTGTGGCCTGA 972
Db      1    AGCAATCAGTAACGGGGCTTACCGTAGCGCGCCCAAGCCGTCGATCGCTGTGGCCTGA 60
QY      973  TCCGTTGGCTGTTAATGTTGAACCCCAACCCGTCGGGCCACACCTGGAGGATGTAGCCAGTC 1032
Db      61   TCCGTTGGCTGTTAATGTTGAACCCCAACCCGTCGGGCCACACCTGGAGGATGTAGCCAGTC 120
QY      1033 ATTGTTGGGTCAACTGGGTTTACACCCGAGTTCGGGGAACAGGAAGCCCTCGTGAGG 1092
Db      121  ATTGTTGGGTCAACTGGGTTTACACCCGAGTTCGGGGAACAGGAAGCCCTCGTGAGG 180
QY      1093 GTGGGCAACCTAGTGTGCTTTGGCGGGCTTCATGGCGGACTGGTACGTCGCTCCT 1152
Db      181  GTGGGCAACCTAGTGTGCTTTGGCGGGCTTCATGGCGGACTGGTACGTCGCTCCT 240
QY      1153 CGGCCCCCTCTCTGGAGAATGGAGCCAAAGTGTGCAGCTTCTTCAAGCAGCAGCTGCCGG 1212
Db      241  CGGCCCCCTCTCTGGAGAATGGAGCCAAAGTGTGCAGCTTCTTCAAGCAGCAGCTGCCGG 300
QY      1213 GAGTGGAAAGCACTGTACCTGGGCTGGAGCGGCAACATTTCTTTAAGAAAGTCCCGAAAGG 1272
Db      301  GAGTGGAAAGCACTGTACCTGGGCTGGAGCGGCAACATTTCTTTAAGAAAGTCCCGAAAGG 360
QY      1273 AGAATCACATGGCTCAAAATCTGCAGAGTGAACCGGCTGAGGATACCTCTTCTCGCCCTG 1332
Db      361  AGAATCACATGGCTCAAAATCTGCAGAGTGAACCGGCTGAGGATACCTCTTCTCGCCCTG 420
QY      1333 GCAAGAGCAGCCTTAAGCTTCCGAAAGGCAATTTCTCAAGAAAAGTCTCTTACCTCGTCAG 1392
Db      421  GCAAGAGCAGCCTTAAGCTTCCGAAAGGCAATTTCTCAAGAAAAGTCTCTTACCTCGTCAG 480
QY      1393 GGGAGTACAGGAGGACCTCTCAGGAATCTCAGACCGGTGCTGATATCTCAGGGCAGCCTG 1452
Db      481  GGGAGTACAGGAGGACCTCTCAGGAATCTCAGACCGGTGCTGATATCTCAGGGCAGCCTG 540
QY      1453 TCCTCTCTGTATCCCTGCTCCCAAGGAAGGCAATCTCTTAAGAGTCTCCAGACGCGTGAAT 1512
Db      541  TCCTCTCTGTATCCCTGCTCCCAAGGAAGGCAATCTCTTAAGAGTCTCCAGACGCGTGAAT 600
QY      1513 CTGGTTACTACTCTCTCCAGAGCCAGCGAGTCTGGGGAACTCTTTAGACGCCAGTGTATG 1572

```



Db 601 CTGTTTACTACTCTCTCCAGAGCCCGAGTCTGGGAACTCTTAGACGCCAGTGATG 660  
 1573 TGTTTGTGAGTG 1584  
 QY  
 Db 661 TGTTTGTGAGTG 672  
 1573 TGTTTGTGAGTG 1584  
 661 TGTTTGTGAGTG 672

RESULT 14  
 CO041415/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

CO041415 716 bp mRNA linear EST 10-JUN-2004  
 UT-M-FW0-cbz-i-16-0-UI.s3 NIH\_BMAP\_FW0 Mus musculus cDNA clone  
 IMAGE:6816569 3', mRNA sequence.  
 CO041415  
 CO041415.1 GI:48581955  
 EST.  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus  
 1 (bases 1 to 716)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-x@mail.nih.gov  
 Tissue Procurement: Dr. James Lin, University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
 http://genome.uiowa.edu/distribution/mousefl.html  
 This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)

The following repetitive elements were found in this cDNA  
 sequence: 1-75, >POLY A#Simple\_repeat (matched complement) 297-361,  
 >(CAAA)N#simple\_repeat  
 Seq primer: pYX-3  
 POLYA=Yes.

FEATURES  
 source  
 Location/Qualifiers  
 1. 716  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6816569"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="NIH\_BMAP\_FW0"  
 /note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;  
 Site 2: Not I; The library was constructed according  
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dr  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated  
 with EcoR I adaptor, digested with NotI and then cloned  
 directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is AGCGAGACAG. This library was created for the University  
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
 Developing Mouse Nervous System', supported by National  
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
 program coordinator.  
 TAG\_SEQ=None found"

ORIGIN  
 Query Match 23.0%; Score 667.4; DB 7; Length 716;  
 Best Local Similarity 99.9%; Pred.No.3.4e-146;  
 Matches 668; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 2230 CTCCTTTGTGCGATGTGTCACCCATGACATTTCCACACCTGTCTTGCTCACCTT 2289

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

Plate: IRBD13 row: a column: 01  
High quality sequence start: 95  
High quality sequence stop: 347.

## FEATURES

source

1. .986

/organism="Mus musculus"

/mol type="mRNA"

```
/db xref="taxon:10090"
```

```
/clone="IMAGE:6431112"
```

```
/etc/hosts: 192.168.0.45
```

[illegible]

```

/clone_id= NIH_MGC_137
/notes="Organ: pancreas; Vector: rspbpm1; site 1: salt."

```

Note: Organ: pancreas; vector: pSPK11; site: 1; Site 2: NotI; Library consists of a pool of clones rearranged from the following libraries: Melton normalized mixed mouse pancreas 1 N1-MMS1, Amplified Melton mouse islets 1 MIS1-A, and Kaestner ng3 wt. Clones rearranged in the laboratory of K. Kaestner (University of Pennsylvania). Note: this is a NIH MGC Library."

## ORIGIN

Query Match 22.8%; Score 663; DB 7; Length 986;

Best Local Similarity 96.2%; Pred. No. 3.9e-145;

2000 Local Similarity	50.2%	FREQ: NO: 3:9E-149;
Matches 709; Conservative	0: Mismatches	24: Indels
	4: Gang	3:

QV 2166 TGTGGAAAGGAA TGGGAACCCCTTGCTGCCCGAG-TGTTATAGTGGGGTGGCGTGAAGGTG 2224

Dh 738 TGGAAAAGGAA TGGGAA CCCCTGCTGCCCGA CTTCTTTATATCTGCGCTCCCTCA ACCCTC 638

**0v**      2225 CATTACCTACTTATGTCGC - ATCGAGTTGTCCGCCCATCCTCACTAACCACGGCCGCGC  
2862

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

[illegible][illegible]

\_\_\_\_\_

[illegible][illegible][illegible]

QY Z403 GGACCAAGATGCCCCCTGGACCCCACTTAGATCTCAGAGACCTTGAACCTTGAAGCTGTTC 2522

438 GGACCCAGATGCCCCCTGGACCCCATCTTAGATCTCAGAGACTTGACCCCTGAAGCTGTCTCT 379

QY 2523 TAGTACCCAGATGTGGATGGATGCTCTGTTCTCAGGCCAACGGGACCTAGAAATGTGCTG 2582

Db 378 TAGTACCCAGATGTGGATGGATGCTCTGTTTCTCAGGCCACCGGCCCTAGAAATGTGCTG 319

Qy 2583 ACTATTATTTTGTGATTCCTCTGTTTTTGG-TTTTTGTTGTTGTTGT 2641

Db 318 ACTATTATTATTTTGTGATTCTCAGTCTCTGTTTTTTGGTTTTTTGTTGTTGTTGGT 259

QY 2642 TTTTGTTTTAAAGTGAATTTTGCTGCTTCAATAATGTGAATGCTGTGTTCTGGGAACT 2701

Db 258 TTTTGTTTTAAAGTGAAATTTTGCTGCTTCAATAATGTGAATGCTGTGTCTCTGGGAACT 199

QY 2702 CCACTGTGCCACTGAAGTTTATGTACAGAGAAGTATTGGCAATGATGTCCCTCTATTCA 2761

db  
198 CCAC TGTGCCACTGAAGTTTATGTACAGAGAAGTATTTGGCAATGATGTCCCTTTATTCA 139

[illegible][illegible]

Db  
78 TCCCTTCACACCCGAAGGCTGCCACCTCCCTCATCTTCATCTGTGGCAAAAAAAAAA 19

Qy 2882 AAAAAAAAAAAAAA 2898

Db 18 AAAAAAAAAAAAAAAAAA 2

Search completed: January 24, 2005, 20:17:42

Job time : 8923.77 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 06:22:58 ; Search time 159.178 Seconds  
(without alignments)  
1422.048 Million cell updates/sec

Title: US-09-980-464-11

Perfect score: 3293

Sequence: 1 MESVALLQRPQAPSASALA.....DCQSVTAAYRALGICSKLS 631

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 200273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_23Sep04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3293	100.0	631	4	AAB50056 Murine Ly
2	3068.5	93.2	630	5	Aae19885 Rat SNF1/
3	2824.5	85.8	628	4	Aam93360 Human pol
4	2824.5	85.8	628	4	Abu53319 Human cel
5	2824.5	85.8	628	5	Abp69116 Human pol
6	2824.5	85.8	628	5	Aau79652 Human pro
7	2824.5	85.8	628	7	Adf76965 Novel hum
8	2824.5	85.8	628	8	Adl30886 Human pro
9	2824.5	85.8	628	8	Adl25362 Human SNA
10	2824.5	85.8	628	8	Adl14161 Novel hum
11	2824.5	85.8	628	8	Ado20172 Human pro
12	2824.5	85.8	672	8	Adj96620 Human cal
13	2800	85.0	629	4	Aab65632 Novel pro
14	2800	85.0	629	8	Adi29239 Human MAR
15	2720.5	82.6	594	5	Aae16266 Human kin
16	2590	78.7	611	6	Abp96085 Human pro
17	2416.5	73.4	534	4	Aab71959 Human TGF
18	1781.5	54.1	406	5	Abp51382 Human MDD
19	1781.5	54.1	406	5	Abp51482 Human MDD
20	1699.5	51.6	661	4	Aae07847 Human pro
21	1699.5	51.6	661	4	Aae07846 Human pro
22	1699.5	51.6	661	7	Ade38421 Human pro
23	1699.5	51.6	661	8	Adj75331 Marker ge
24	1699.5	51.6	661	8	Adl25353 Human ARK
25	1699.5	51.6	661	8	Adq19734 Human sof

#### ALIGNMENTS

##### RESULT 1

AAB50056  
ID AAB50056 standard; protein; 631 AA.

XX AAB50056;

DT 19-MAR-2001 (first entry)

DE Murine Lymph node Stromal cell kinase 1.

XX Murine; Lymph node Stromal cell kinase; MUSK-1; autoimmune disorder;

KW wound healing; periodontal disease; inflammatory disease; tumour;

KW infection; allergy.

XX Mus musculus.

XX WO200073468-A1.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US014696.

XX 28-MAY-1999; 99US-0136781P.

XX (IMMV ) IMMUNEX CORP.

XX Bird TA, Virca GD, Martin U, Anderson DM;

XX WPI; 2001-061546/07.

XX N-PSDB; AAC90433.

XX Novel murine and human kinase nucleic acids useful for treating inflammations, infections, tumors, allergies, autoimmune diseases, and for stimulating or suppressing immune responses.

XX Claim 10; Page 94-96; 106pp; English.

XX The present sequence is Murine Lymph node Stromal cell kinase 1 (MUSK-1). This protein is useful for treating a variety of disorders listed in the disclosure of the specification, including autoimmune disorders, allergic reactions, myeloid or lymphoid cell deficiencies, wound healing, and tissue repair and replacement, burns, incisions and ulcers, periodontal disease, inflammatory diseases, tumours and bacterial, viral or fungal infection

XX Sequence 631 AA;

Query Match 100.0%; Score 3293; DB 4; Length 631;

Best Local Similarity 100.0%; Pred. No. 9.5e-281;  
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESVALLORPQSPASALASASAPLADGLIKSPKPLMKQAVKRRHHKINLHRHYEFL 60  
Db 1 MESVALLORPQSPASALASASAPLADGLIKSPKPLMKQAVKRRHHKINLHRHYEFL 60

Qy 61 ETLGKGTGKVKKARESSGRVVAIKSIRKDKIKDEQDLHTRREIEMSSLNHPHIIAH 120  
Db 61 ETLGKGTGKVKKARESSGRVVAIKSIRKDKIKDEQDLHTRREIEMSSLNHPHIIAH 120

Qy 121 EVFENSKIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHQNGIVHRDL 180  
Db 121 EVFENSKIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHQNGIVHRDL 180

Qy 181 KLENILLDANGNIKIADFGCLNLYHKGFQTFGSPLYASPEIVNGKPYVGPVEDSWSL 240  
Db 181 KLENILLDANGNIKIADFGCLNLYHKGFQTFGSPLYASPEIVNGKPYVGPVEDSWSL 240

Qy 241 GVLLYILVHGTMPDGDQDKTLVKQISNGAYREPPKPSDACGLIRWLLMVNTPTRATLED 300  
Db 241 GVLLYILVHGTMPDGDQDKTLVKQISNGAYREPPKPSDACGLIRWLLMVNTPTRATLED 300

Qy 301 VASHWVNWGTYTGVGEALREGHPSGDFGRASMDWLRSSRPILLENGAKVCSPFKQ 360  
Db 301 VASHWVNWGTYTGVGEALREGHPSGDFGRASMDWLRSSRPILLENGAKVCSPFKQ 360

Qy 361 HVPGGSTVPGLEQHSLSKKRKNEMDAQNLOQDPAEDTSSRPKSSILKPKGILKKSS 420  
Db 361 HVPGGSTVPGLEQHSLSKKRKNEMDAQNLOQDPAEDTSSRPKSSILKPKGILKKSS 420

Qy 421 TSSEVQDPELRPVPPTGQPVAVSLLPRKGLKKSRORESGYTSSPEPSSGELL 480  
Db 421 TSSEVQDPELRPVPPTGQPVAVSLLPRKGLKKSRORESGYTSSPEPSSGELL 480

Qy 481 ASDVFSVSDPVEQKSPQASGLLHHRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540  
Db 481 ASDVFSVSDPVEQKSPQASGLLHHRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540

Qy 541 ARPSRPSGAVSEDISSFDQDLPERLPETPLRGCVSDNLRGLRQPPSEGLKRWQ 600  
Db 541 ARPSRPSGAVSEDISSFDQDLPERLPETPLRGCVSDNLRGLRQPPSEGLKRWQ 600

Qy 601 ESLGDSFSLTDCQEVTAAYRQALGICSKLS 631  
Db 601 ESLGDSFSLTDCQEVTAAYRQALGICSKLS 631

## RESULT 2

AAE19885  
ID AAE19885 standard; protein; 630 AA.  
XX  
AC AAE19885;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Rat SNF1/AMPK-Related Kinase (SNARK) protein.  
XX  
KW Rat; SNF1/AMPK-Related Kinase; SNARK; enzyme; stress response; diabetes;  
KW glucose deprivation; lipid metabolism; therapy; lipoprotein disorder;  
KW hyperglycaemic; drug screening; hypoglycaemia.  
XX  
OS Rattus sp.  
XX  
FH Key Location/Qualifiers  
FT Binding-site 63..89  
FT Domain /note= "Protein kinase ATP-binding region signature"  
FT Domain 137..140  
FT Active-site /note= "Serine/threonine kinase catalytic domain"  
FT Active-site 175..187  
FT Domain /note= "Serine/threonine protein kinase active-site"  
FT Domain 297..300  
FT /note= "Serine/threonine kinase catalytic domain"

FT Domain 335..338  
FT /note= "Serine/threonine kinase catalytic domain"  
FT Domain 381..384  
FT /note= "Serine/threonine kinase catalytic domain"  
FT Domain 422..425  
FT /note= "Serine/threonine kinase catalytic domain"  
FT Domain 468..471  
FT /note= "Serine/threonine kinase catalytic domain"  
FT Domain 517..520  
FT /note= "Serine/threonine kinase catalytic domain"  
FT Domain 601..604  
FT /note= "Serine/threonine kinase catalytic domain"  
FT Domain 608..611  
FT /note= "Serine/threonine kinase catalytic domain"

W0200212456-A2.

14-FEB-2002.

02-AUG-2001; 2001WO-CA001109.

03-AUG-2000; 2000US-022250P.

12-MAR-2001; 2001US-0274613P.

28-MAR-2001; 2001CA-02340783.

(ONEO-) 1149336 ONTARIO INC.

Drucker DJ, Rosen CF, Lefebvre DL;

WPI; 2002-241747/29.

N-PSDB; AAD31710.

AMPK (AMP-activated protein kinase)-related kinase, designated SNARK polypeptides and polynucleotides, useful for treating or preventing diabetes, or other disorders of lipoprotein production leading to increased levels of cholesterol.

Claim 1; Fig 2; 94pp; English.

The invention relates to an AMPK (AMP-activated protein kinase)-related kinase, designated SNARK polypeptides and polynucleotides. SNARK (SNIF/AMP-activated protein kinase) is involved in stress response to glucose deprivation. The polynucleotides are useful for expressing SNARK protein in isolated form or as a protein conjugate. Activation of SNARK stimulates liver CPT-1 thus enhances lipid metabolism in liver cells and in other cell types such as heart and skeletal muscles, as well as increases GLUT-4 and glycogen in muscle. Activation of SNARK is predicted to have insulin-like effects that would enhance the disposal of glucose into muscle and reduce plasma glucose for the treatment of diabetes and some type of disorders of lipoprotein production leading to increased levels of cholesterol or triglycerides. SNARK or its variants may be administered to a subject to treat or prevent a disease associated with decreased expression of SNARK, such as diabetes. SNARK antibodies are used to modulate SNARK activity either in vivo for therapeutic purposes, or in vitro, for drug screening and related investigational purposes. SNARK antagonists may be administered to increase fuel production. decrease glucose uptake and increase levels of blood glucose in a patient suffering from hypoglycaemia. The present sequence is rat SNARK protein

Sequence 630 AA;

Query Match 93.2%; Score 3068.5; DB 5; Length 630;  
Best Local Similarity 94.0%; Pred. No. 5.9e-261;  
Matches 593; Conservative 12; Mismatches 25; Indels 1; Gaps 1;

Qy 1 MESVALLORPQSPASALASAPLADGLIKSPKPLMKQAVKRRHHKINLHRHYEFL 60  
Db 1 MESVALLORPQSPASALASAPLADGLIKSPKPLMKQAVKRRHHKINLHRHYEFL 60

Qy 61 ETLGKGTGKVKKARESSGRVVAIKSIRKDKIKDEQDLHTRREIEMSSLNHPHIIAH 120  
Db 61 ETLGKGTGKVKKARESSGRVVAIKSIRKDKIKDEQDLHTRREIEMSSLNHPHIIAH 120

```
QY 121 EVFENSKIIVMEYASRGDLVYISERPRLSERDARHFFRQIVSALHYCHONGIVHRDL 180
Db 121 EVFENSKIIVMEYASRGDLVYISERPRLSERDARHFFRQIVSALHYCHONGIVHRDL 180
QY 181 KLENILLDANGNIKIADPGLSNLHKGFQTCGSPLYASPELVNKPYPVGPVDSWSL 240
Db 181 KLENILLDANGNIKIADPGLSNLHKGFQTCGSPLYASPELVNKPYPVGPVDSWSL 240
QY 241 GVLlyILVHGTMPFDGQDHKTIVKQISNGAYREPPKPSDACGLIRWLLMNPTRATLED 300
Db 241 GVLlyILVHGTMPFDGQDHKTIVKQISNGAYREPPKPSDACGLIRWLLMNPTRATLED 300
QY 301 VASHWWNMGYTTGVGQEQALRECGHPSGDFGRASMDWLRSSRPLLENGAKVCSFFKQ 360
Db 301 VASHWWNMGYTTGVGQEQALRECGHPSGDFGRASMDWLRSSRPLLENGAKVCSFFKQ 360
QY 361 HVPGGGTVPLRGHSLKSRKENDMAQNLQDPAEDTSRPGKSLKLPKGLKKSS 420
Db 361 HVPGGGTVPLRGHSLKSRKENDMAQNLQDPAEDTSRPGKSLKLPKGLKKSS 420
QY 421 TSSGEVQEDPQELRPVPTDQGPVPAVSLPRKGLKKSQRRESGYSSPEPSESGELLD 480
Db 421 TSSGEVQEDPQELRPVPTDQGPVPAVSLPRKGLKKSQRRESGYSSPEPSESGELLD 480
QY 481 ASDVFGSDPVEQKSPQASGLLLHHRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540
Db 481 AGDVFGSDPVEQKSPQASGL--RLHHRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540
QY 541 ARSRPFGAVSEDSILSESFDQDLPERLPETPLRCVSDNLRGLFQEPSEGLKRWQ 600
Db 541 ARSRPFGAVSEDSILSESFDQDLPERLPETPLRCVSDNLRGLFQEPSEGLKRWQ 600
QY 601 ESLGDSFSLTDCQEVTAAYRQALGICSKLS 631
Db 601 ESLGDSFSLTDCQEVTAAYRQALGICSKLS 630

RESULT 3
AAM93360
ID AAM93360 standard; protein; 628 AA.
XX
AC AAM93360;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide, SEQ ID NO: 2919.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
XX
PR 11-JAN-2000; 2000JP-0018774.
XX
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
DR WPI; 2001-524255/58.
XX
DR N-PSDB; AAK94280.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 8; SEQ ID NO 2919; 1380pp + Sequence Listing; English.
```

```
XX The invention relates to primers for synthesizing full length cDNA
CC clones, 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesized by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a polypeptide encoded by a full length
CC human cDNA of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in CD-ROM
CC format directly from EPO
XX
SQ Sequence 628 AA;
```

```
Query Match 85.8%; Score 2824.5; DB 4; Length 628;
Best Local Similarity 85.7%; Pred. No. 1.9e-239;
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;
QY 1 MESVALLQRPSPQASASALASERPLADGLIKSPKPKMKKQAVKRHHKHLRHRVEFL 60
Db 1 MESLVPARRSGPTFS-----AAELARPLAEGLIKSPKPKMKKQAVKRHHKHLRHRVEFL 56
QY 61 ETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEQDLHRRREIEIMSSLNHPHIIAH 120
Db 57 ETLGKGTGKVKKARESSGRLVAIKSIRKDKIKDEQDLHRRREIEIMSSLNHPHIIAH 116
QY 121 EVFENSKIIVMEYASRGDLVYISERPRLSERDARHFFRQIVSALHYCHONGIVHRDL 180
Db 117 EVFENSKIIVMEYASRGDLVYISERQQLSREARHFFRQIVSAVHYCHQNVRVHRDL 176
QY 181 KLENILLDANGNIKIADPGLSNLHKGFQTCGSPLYASPELVNKPYPVGPVDSWSL 240
Db 177 KLENILLDANGNIKIADPGLSNLHKGFQTCGSPLYASPELVNKPYPVGPVDSWSL 236
QY 241 GVLlyILVHGTMPFDGQDHKTIVKQISNGAYREPPKPSDACGLIRWLLMNPTRATLED 300
Db 237 GVLlyILVHGTMPFDGQDHKTIVKQISNGAYREPPKPSDACGLIRWLLMNPTRATLED 296
QY 301 VASHWWNMGYTTGVGQEQALRECGHPSGDFGRASMDWLRSSRPLLENGAKVCSFFKQ 360
Db 297 VASHWWNMGYATRVGQEAPHEGGHPSGDSASASMDWLRSSRPLLENGAKVCSFFKQ 356
QY 361 HVPGGGTVPLRGHSLKSRKENDMAQNLQDPAEDTSRPGKSLKLPKGLKKSS 420
Db 357 HAFGGGTTPLRGHSLKSRKENDMAQNLQDPAEDTSRPGKSLKLPKGLKKSS 416
QY 421 TSSGEVQEDPQELRPVPTDQGPVPAVSLPRKGLKKSQRRESGYSSPEPSESGELLD 480
Db 417 ASAEVQEDPPELSPITPASPGQAAP---LLPKKGLKPKQRESGYSSPEPSESGELLD 473
QY 481 ASDVFGSDPVEQKSPQASGLLLHHRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540
Db 474 AGDVFGSDPVEQKSPQASGLLLHHRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 533
QY 541 ARSRPFGAVSEDSILSESFDQDLPERLPETPLRCVSDNLRGLFQEPSESGELLD 596
Db 534 ARSRPFGAVSEDSILSESFDQDLPERLPETPLRCVSDNLRGLFQEPSESGELLD 593
QY 597 RWMQESLGDSFSLTDCQEVTAAYRQALGICSKLS 631
Db 594 RWRQDPLGDSFSLTDCQEVTAAYRQALGICSKLS 628
```

```
RESULT 4
ABU53319
ID ABU53319 standard; protein; 628 AA.
XX
AC ABU53319;
XX
DT 14-APR-2003 (first entry)
XX
```

Human cell cycle-associated protein from DKFZphtes3\_7j3.

Human; gene therapy; vaccine; disease treatment; detection.

Homo sapiens.

WO200112659-A2.

22-FEB-2001.

18-AUG-2000; 2000WO-IB001496.

18-AUG-1999; 99US-0149499P.

28-SEP-1999; 99US-0156503P.

(GEHU-) GERMAN HUMAN GENOME PROJECT.

Wiemann S;

WPI; 2001-327840/34.

N-PSDB; ABX71420.

Nucleic acids having the sequences of clones isolated from libraries of different human tissues, useful in recombinant DNA methodologies.

Claim 21; Page 943; 1095pp; English.

This invention describes novel polynucleotides and polypeptides isolated from human cDNA libraries which can be used for gene therapy or in vaccines. The polynucleotides of the invention and antibodies encoded by them may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate polypeptide expression. The products of the invention may also be used to identify modulators of expression and activity and to down regulate expression and activity. The antibodies of the invention may also be used as diagnostic agents for detecting the presence of polypeptides in samples. This sequence represents a polypeptide described in the disclosure of the invention

Sequence 628 AA;

Query Match 85.8%; Score 2824.5; DB 4; Length 628;

Best Local Similarity 85.7%; Pred. No. 1.9e-239;

Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;

QY 1 MESVALLQRPQAPASALASESARPLADGLIKSPKLMKQAVKRRHHKHLRHYEFL 60

DB 1 MESLVFARRSGTPS----AAELARPLAELIKSPKLMKQAVKRRHHKHLRHYEFL 56

QY 61 ETLLGKTYGVKKARESSGRGLVAIKSRKDKIODEQLLHRRRIEIMSSLNHPHIIAH 120

DB 57 ETLLGKTYGVKKARESSGRGLVAIKSRKDKIODEQLLHRRRIEIMSSLNHPHIIAH 116

QY 121 EVFENSSTVIMVYASRGDLVYISERPLSERDARHFRQIVSALHYCHONGIVHRDL 180

DB 117 EVFENSSTVIMVYASRGDLVYISERQQLSERARHFRQIVSALHYCHONRVHRDL 176

QY 181 KLENILDANGNIKIADFGSLNLYHKGKFLQTCGSPLYASPELVNGKPYGVPEVDSWSL 240

DB 177 KLENILDANGNIKIADFGSLNLYHKGKFLQTCGSPLYASPELVNGKPYGVPEVDSWSL 236

QY 241 GVLLYILVHGTMPPDGQDHTLVKQISNGAYREPPKPSDACGLIRLLWMNPTFRATLED 300

DB 237 GVLLYILVHGTMPPDGQDHTLVKQISNGAYREPPKPSDACGLIRLLWMNPTFRATLED 296

QY 301 VASHWWNWYTTGVGQEQALREGCHPSGDFGRASMDWLRSSRPLLENGAKVCSPFKQ 360

DB 297 VASHWWNWYTTGVGQEQALREGCHPSGDFGRASMDWLRSSRPLLENGAKVCSPFKQ 356

QY 361 HVPGGGTVPLERQHSLSKSRKENDMAQNLQGDPAEDTSRPGKSLKLPKGLKKKS 420

DB 357 HAPGGGTTVPLERQHSLSKSRKENDMAQNLQHSADTDTHRPGKSLKLPKGLKKKS 416

QY 421 TSSEVOEDPQELRPVPTTFQGPVPAVSLPRKGIKKSRQREGSGYSSPEPSGELL 480

Db 417 ASAEVQEDPELSPASPQQAAP---LLPKKGIKKRQREGSGYSSPEPSGELL 473

Qy 481 ASDVFSVSGDPVEQKSPQASGLLLHRKGIKLNKGFSTALEGTTPTFGSLDQLASSHPA 540

Db 474 AGDVFSVSGDPKEQKPPQASGLLLHRKGIKLNKGFSTALEAAPTTFGSLDELAPRPL 533

Qy 541 ARPSRPSGAVSEDSILSSESFQDLDLPERLPETPLRGCVSDNLRGLQEPSPSG---LK 596

Db 534 ARASRPSGAVSEDSILSSESFQDLDLPERLPETPLRGCVSDNLTGLEPSPSGCLR 593

Qy 597 RWOESLGDSFSLTDCQEVTAAYRQALGICSKLS 631

Db 594 RWRQDPLGDSFSLTDCQEVTAAYRQALRVCSKLT 628

RESULT 5

ABP69116

ID ABP69116 standard; protein; 628 AA.

XX AC ABP69116;

XX DT 20-JAN-2003 (first entry)

XX DE Human polypeptide SEQ ID NO 1163.

XX KW Human; genome mapping; gene therapy; food supplement; virus; fungus; cell-proliferative disorder; neurodegenerative disease; bacterial; Parkinson's disease; Alzheimer's disease; autoimmune disease; multiple sclerosis; diabetes; genetic disorder; wound; burn; infection; arthritis; cytostatic; immunomodulator; neotropic; neuroprotective; antiparkinsonian; antidiabetic; immunosuppressive; dermatological; haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide; antiarthritic.

XX OS Homo sapiens.

XX PN WO200270539-A2.

XX PD 12-SEP-2002.

XX PF 05-MAR-2002; 2002WO-US005095.

XX PR 05-MAR-2001; 2001US-00799451.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren P;

XX PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;

XX PI Wehrman T, Wang J, Wang D, Drmanac RT;

XX DR WPI; 2002-759812/82.

XX DR N-PSDB; ABZ11333.

XX PT New polynucleotides comprising sequences assembled from expressed sequence tags (ESTs), useful for treating cell-proliferative, neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet or coagulation disorders.

PS Claim 9; SEQ ID NO 1163; 1012pp + Sequence Listing; English.

XX CC The invention relates to an isolated polynucleotide (1) comprising a nucleotide sequence selected from any of 948 sequences (ABZ11119-ABZ12066) or their mature protein coding portion, active domain coding protein or complementary sequences. The polynucleotides are useful for identifying expressed genes or for physical mapping of human genome. The encoded polypeptides (ABP6902-ABP69849) are useful as molecular weight markers, as a food supplement, for generating antibodies, in medical imaging, screening and diagnostic assays and for treating cell-proliferative disorders (cancer), neurodegenerative diseases (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid disorders, platelet or coagulation disorders, wound, burns, incision, ulcers, liver





297 VASHWVNWGYATRVGEQEPHGGHSGDSARASMDWLRSSRPILLENGAKVCSFFKQ 356  
 361 HVPGGSTVPGLEHQHSLKKRKENDMAQNLQDPAEDTSRPOKSSKLKPKGILKKSS 420  
 357 HAPGGSTTPGLERQHSLKKRKENDMAQSLHSDTADTAHRPGKSNLKLKPKGILKKVS 416  
 421 TSSEVOEDPOELRPVDPDPGPVAVSLPRKGLKKSRORESGYSSPSPSSGELL 480  
 417 ASAGVQEDPELSPFASPQOAP---LLPKGLKKPRORESGYSSPSPSSGELL 473  
 481 ASDVFGSDPVVEQKSPQASGLLHRKGLKNGKFSRTALRGTPSTFGSLDOLASSHPA 540  
 474 AGDVFGSDPVVEQKSPQASGLLHRKGLKNGKFSRTALEAAPTTFGSLDELAPRPL 533  
 541 ARPSRPSGAVSEDSILSSESDQLDLPRLPETPLRGCVSDNLRGLSEQPSSEG---LK 596  
 534 ARASRPSGAVSEDSILSSESDQLDLPRLPETPLRGCVSDNLTGLEEPPSGGSCLR 593  
 597 RWMQESLGDSFCFLTDQCEVTAAYRQALGICSKLS 631  
 594 RWRQDPLGDSFCFLTDQCEVTATYRQALRVCSKLT 628

RESULT 7  
 ADF76965  
 ID ADF76965 standard; protein; 628 AA.  
 AC ADF76965;  
 XX  
 XX 26-FEB-2004 (first entry)  
 XX  
 XX Novel human secreted and transmembrane protein SeqID 640.  
 XX human; PRO; membrane bound protein; membrane bound receptor;  
 KW cell proliferation; cell migration; cell differentiation;  
 KW mitogenic factor; survival factor; cytotoxic factor;  
 KW differentiation factor; neuropeptide; hormone; cell receptor;  
 KW receptor-ligand interaction; cytostatic; chondrocyte; tumour.  
 XX  
 XX Homo sapiens.  
 XX  
 XX WO2003072035-A2.  
 XX  
 XX 04-SEP-2003.  
 XX  
 XX 21-FEB-2003; 2003WO-US005241.  
 XX  
 XX 22-FEB-2002; 2002US-0359461P.  
 XX  
 XX (GETH ) GENENTECH INC.  
 XX  
 XX Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;  
 XX Williams FM, Wood WI, Wu TD;  
 XX WPI; 2003-721702/68.  
 XX  
 XX N-PSDB; ADF76964.  
 XX  
 XX New PRO polypeptides, useful for diagnosing and treating an immune  
 XX related disorder, e.g. systemic lupus erythematosus, rheumatoid  
 XX arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or  
 XX diabetes mellitus.  
 XX  
 XX Claim 10; SEQ ID NO 640; 918pp; English.

cell receptors or membrane bound proteins. These membrane bound proteins  
 and receptors may be of use as pharmaceutical and diagnostic agents, such  
 as in the blocking of receptor-ligand interactions. The current invention  
 provides the amino acid sequences of novel human membrane bound receptors  
 and proteins, along with the cDNA sequences encoding them. The novel  
 proteins of the invention may have cytostatic activities through the  
 stimulation of chondrocytes. The nucleic acids of the invention may be  
 useful for the manufacture of a medicament for diagnosing or treating a  
 tumour in a mammal. In addition, they may be useful for measuring or  
 detecting the expression of a tumour associated gene. The present  
 sequence is the amino acid sequence of a human PRO protein of the  
 invention.

XX  
 XX  
 XX Sequence 628 AA;  
 SQ

Query Match 85.8%; Score 2824.5; DB 7; Length 628;  
 Best Local Similarity 85.7%; Pred. No. 1.9e-239;  
 Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;

QY 1 MESVALLORPSQASALASESARPLADGLIKSPKLMKKQAVKRHHKHLRHYEFL 60  
 DB 1 MESLVFARRSGPTPS-----AAELARPLAELGLIKSPKLMKKQAVKRHHKHLRHYEFL 56  
 QY 61 ETILGKTYGVKVKARESSGRLVAIKSRDKIKDEQDLLHRRREIEIMSSLNHPHIIAIIH 120  
 DB 57 ETILGKTYGVKVKARESSGRLVAIKSRDKIKDEQDLLHRRREIEIMSSLNHPHIIAIIH 116  
 QY 121 EVFENGSKIIVIMVEYASRGDLYIISERPLSRDARHFFRQIVSALHYCHONGIVHRDL 180  
 DB 117 EVFENGSKIIVIMVEYASRGDLYIISERQQLSREARHFFRQIVSAVHYCHONRVHRDL 176-  
 QY 181 KLENILLDANGNIKIADFGLSNLYHKGFLOTGSGPLIYASPEIVNGKPYTGPVEVDSWSL 240  
 DB 177 KLENILLDANGNIKIADFGLSNLYHQGKFLQTTGSGPLIYASPEIVNGKPYTGPVEVDSWSL 236  
 QY 241 GVLLYILVHGTMPFDGQDHKTLVKQISNGAYRPPKPSDACGLIRWLLMWNPTRRATLED 300  
 DB 237 GVLLYILVHGTMPFDGHDHKLILVKQISNGAYRPPKPSDACGLIRWLLMWNPTRRATLED 296  
 QY 301 VASHWVNWGYTTGVGEQEAALREGGHPGSDFGASMDWLRSSRPILLENGAKVCSFFKQ 360  
 DB 297 VASHWVNWGYATRVGEQEPHGGHSGDSARASMDWLRSSRPILLENGAKVCSFFKQ 356  
 QY 361 HVPGGSTVPGLEHQHSLKKRKENDMAQNLQDPAEDTSRPOKSSKLKPKGILKKSS 420  
 DB 357 HAPGGSTTPGLERQHSLKKRKENDMAQSLHSDTADTAHRPGKSNLKLKPKGILKKVS 416  
 QY 421 TSSEVOEDPOELRPVDPDPGPVAVSLPRKGLKKSRORESGYSSPSPSSGELL 480  
 DB 417 ASAGVQEDPELSPFASPQOAP---LLPKGLKKPRORESGYSSPSPSSGELL 473  
 QY 481 ASDVFGSDPVVEQKSPQASGLLHRKGLKNGKFSRTALEGTPSTFGSLDOLASSHPA 540  
 DB 474 AGDVFGSDPVVEQKSPQASGLLHRKGLKNGKFSRTALEAAPTTFGSLDELAPRPL 533  
 QY 541 ARPSRPSGAVSEDSILSSESDQLDLPRLPETPLRGCVSDNLRGLSEQPSSEG---LK 596  
 DB 534 ARASRPSGAVSEDSILSSESDQLDLPRLPETPLRGCVSDNLTGLEEPPSGGSCLR 593  
 QY 597 RWMQESLGDSFCFLTDQCEVTAAYRQALGICSKLS 631  
 DB 594 RWRQDPLGDSFCFLTDQCEVTATYRQALRVCSKLT 628

RESULT 8  
 ADL30886  
 ID ADL30886 standard; protein; 628 AA.  
 XX  
 XX AC ADL30886;  
 XX  
 XX DT 20-MAY-2004 (first entry)  
 XX  
 XX Human protein encoded by a full length cDNA clone SeqID 2919.

human; medicine; signal transduction; glycoprotein; transcription; oligo-capping method.

Homo sapiens.

EPI396543-A2.

10-MAR-2004.

07-JUL-2000; 2003EP-00025638.

08-JUL-1999; 99JP-00194486.

11-JAN-2000; 2000JP-00118774.

02-MAY-2000; 2000JP-00183865.

07-JUL-2000; 2000EP-00114089.

(REAS-) RES ASSOC BIOTECHNOLOGY.

Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y, Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H; WPI: 2004-204755/20.

N-PSDB; ADL30885.

New oligonucleotide primers (830 cDNAs) useful for synthesizing full length human cDNAs.

Example 1; SEQ ID NO 2919; 1340pp; English.

This invention relates to a novel primers useful for synthesizing full length cDNA molecules that encode human proteins. Specifically, it refers to secretory or membrane proteins that are potential therapeutic agents/target molecules in the field of medicine, and in particular genes encoding proteins that are associated with signal transduction, glycoproteins and transcription. The present invention describes a method for efficiently cloning a full length human cDNA from both the 5' and 3' ends using the oligo-capping method. This polypeptide sequence is a full length human protein of the invention.

Sequence 628 AA;

Query Match 85.8%; Score 2824.5; DB 8; Length 628;  
Best Local Similarity 85.7%; Pred. No. 1.9e-239;  
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;

QY 1 MESVALLQRPQAPASALASESARPLADGLIKSPKLMKKQAVKRHHKHLRHYEFL 60  
DB 1 MESLVFARRSGPTPS-----AAELARPLAEGLIKSPKLMKKQAVKRHHKHLRHYEFL 56

QY 61 ETLCGTYGKVKKARESGLVAIKSRKDKIKDEQDLHTRRIEIMSSLNHPHIIAIIH 120  
DB 57 ETLCGTYGKVKKARESGLVAIKSRKDKIKDEQDLHTRRIEIMSSLNHPHIIAIIH 116

QY 121 EVFENSISKIVMEYASRGDLYDIYSRPRLSERDARHFRQIVSALHYCHQNGIVHRDL 180  
DB 117 EVFENSISKIVMEYASRGDLYDIYSRQQLSERARHFRQIVSAVHYCHQNRVHRDL 176

QY 181 KLENILLDANGNIKIDFGLSNLYHKGFLOTFGSPLYASPEIYVNGKPYVGPVDSWSL 240  
DB 177 KLENILLDANGNIKIDFGLSNLYHOGKFLQTFGSPLYASPEIYVNGKPYTGPVDSWSL 236

QY 241 GVLLYILVHGTMPDGDHKTIVKQISNGAYREPDKSDAGLIRWLLMWNPTRRATLED 300  
DB 237 GVLLYILVHGTMPDGDHKTIVKQISNGAYREPDKSDAGLIRWLLMWNPTRRATLED 296

QY 301 VASHWWNVGTYTGVGQEQALREGCHPSGDFGRASMDLWLRSSRPLLENGAKVCSFFKQ 360  
DB 297 VASHWWNVGTYATRVGQEQAPHEGCHPSGDSARASMDLWLRSSRPLLENGAKVCSFFKQ 356

QY 361 HVPGGSGTVPGLERQHSLSKSRKENDMAQNLQGDPAEDTSSRPKSSLSKLPKGLKKKSS 420  
DB 357 HAPGGSGTTPGLERQHSLSKSRKENDMAQSLHSDTADDTAHRPGKSNLKLPGILKKKVS 416

QY 421 TSSEGEVQEDPOELRPVPTPGQPVPAVSLIPRKILKKSRORESGYSSPPSPSGELLD 480  
DB 417 ASAEVQEDPELSPIPASFGQAP---LLPKGILKPKQRESGYSSPPSPSGELLD 473

QY 481 ASDVFSVSGDPVEQKSPQASGLLLHRKIGILKNGKFSRTALEGTTPTFGSLDQLASSHPA 540  
DB 474 AGDVFSVSGDPKEQKPPQASGLLLHRKIGILKNGKFSOTALELAAPTTFGSLDELAPRPL 533

QY 541 ARSPRPSGAVSEDSILSSSEFQDLDPRLPETPLRCVSDNLRGLQEPSEG----LK 596  
DB 534 ARASRPSGAVSEDSILSSSEFQDLDPRLPETPLRCVSDNLTGLEPSPSGSGCLR 593

QY 597 RWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631  
DB 594 RWRQDPLGDSFSLTDCQEVTAAYRQALRVCSKLT 628

RESULT 9  
ADL25362  
ID ADL25362 standard; protein; 628 AA.  
XX  
AC ADL25362;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Human SNARK, SEQ ID 12.  
XX  
KW Cytostatic; Neuroprotective; Nootropic; Muscular; Gene therapy;  
KW Adenosine Monophosphate activated Protein Kinase-related kinase 5;  
KW AMPK-related kinase 5; ARK5; KIAA0537; stress resistance; tumour;  
KW nervous disorder; muscle disorder; ataxia teleangiectasia; SNARK.  
OS Homo sapiens.  
XX WO2004019994-A1.  
XX 11-MAR-2004.  
XX 19-AUG-2003; 2003WO-JP010435.  
XX 27-AUG-2002; 2002JP-00247761.  
XX (NINA-) JAPAN AGENCY NAT INST HEALTH.  
XX (PHAR-) ORG PHARM SAFETY & RES.  
XX  
PI Esumi H, Suzuki A;  
XX  
DR WPI; 2004-248195/23.  
XX  
PT Agent for imparting stress resistance to cells, comprises DNA encoding  
PT for adenosine monophosphate activated protein kinase-related kinase 5 for  
PT treatment of tumours, muscle and nervous disorders, and ataxia  
PT teleangiectasia.  
XX  
PS Example 1; SEQ ID NO 12; 143pp; Japanese.  
XX  
CC The present invention relates to human Adenosine Monophosphate activated  
CC Protein Kinase (AMPK)-related kinase 5 (ARK5/ KIAA0537) and its coding  
CC sequence, which are useful for imparting stress resistance to cells. ARK5  
CC is useful for treatment and prevention of tumours, nervous disorders,  
CC muscle disorders and ataxia teleangiectasia. The present sequence was  
CC used to illustrate the invention.  
XX  
SQ Sequence 628 AA;

Query Match 85.8%; Score 2824.5; DB 8; Length 628;  
Best Local Similarity 85.7%; Pred. No. 1.9e-239;  
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;

QY 1 MESVALLQRPQAPASALASESARPLADGLIKSPKLMKKQAVKRHHKHLRHYEFL 60  
DB 1 MESLVFARRSGPTPS-----AAELARPLAEGLIKSPKLMKKQAVKRHHKHLRHYEFL 56

QY 61 ETLGKGTGKVKKARESSGLVAIKSRKDKIKDEQDLHRRREIEMSSLNPHIIAIIH 120  
DB 57 ETLGKGTGKVKKARESSGLVAIKSRKDKIKDEQDLHRRREIEMSSLNPHIIAIIH 116  
QY 121 EVFENSSKIVVMVEYASRGDLXYISRPILSERDARHFFRQIVSALHYCHQNGIVHRDL 180  
DB 117 EVFENSSKIVVMVEYASRGDLXYISERQQLSEREARHFFRQIVSAVHYCHQNRVHRDL 176  
QY 181 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEDVSWSL 240  
DB 177 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYTGPEDVSWSL 236  
QY 241 GVLLYILVHGTMPDGDHKLIVKQISNGAYREPPEKSDACGLIRWLLMNPTRATLED 300  
DB 237 GVLLYILVHGTMPDGDHKLIVKQISNGAYREPPEKSDACGLIRWLLMNPTRATLED 296  
QY 301 VASHWWNVWYTTGVGQEQALREGGHPGDFGRASMDLRRSRPILLENKAKVCSPFKQ 360  
DB 297 VASHWWNVWYATRVGQEQAPHEGGHPGDSARASMDLRRSRPILLENKAKVCSPFKQ 356  
QY 361 HVGCGGSTVPLERQHSLSKSRKENDMAQNLQSDPAEDTSSRPKSSLSKLPKGLKXKSS 420  
DB 357 HAPCGGSTVPLERQHSLSKSRKENDMAQNLQSDPAEDTSSRPKSSLSKLPKGLKXKSS 416  
QY 421 TSSEVEQEDQELRPVPTDTPQVPVAVSLLPRKGLKSRQESGYSSPEPSESGLLD 480  
DB 417 ASAEVQEDPELSPIPASPGQAAP---LPPKGLKSRQESGYSSPEPSESGLLD 473  
QY 481 ASDVVFVSGDPEQKSPQASGLLLHRKGLILKNGKFSRTALEGTTPTFGSLDOLASHPA 540  
DB 474 AGDVVFVSGDPEQKSPQASGLLLHRKGLILKNGKFSRTALEGTTPTFGSLDOLASHPA 533  
QY 541 ARPSRPGAVSEDSILSESPDQLDPLRLPETPLRGVSDNLRGLQEPPESEG---LK 596  
DB 534 ARASRPGAVSEDSILSESPDQLDPLRLPETPLRGVSDNLRGLQEPPESEGSGCLR 593  
QY 597 RWMQESIGDSCFSLTDCQEVTAARQALGICKSL 631  
DB 594 RWRQDPLGDSFSLTDCQEVTAARQALRVCSKLT 628  
RESULT 10  
ADL14161  
XX ID ADL14161 standard; protein; 628 AA.  
XX AC ADL14161;  
XX DT 17-JUN-2004 (first entry)  
XX DE Novel human gene 3700 encoded protein.  
XX KW cytostatic; cardiac; hypotensive; antianginal; osteopathic;  
KW antiarthritic; antirheumatic; neuroprotective; antiinflammatory;  
KW antiparietal; antidiabetic; cardiovascular; viricide; analgesic; CNS;  
KW angiogenesis inhibitor; angiogenesis stimulator; cerebroprotective;  
KW nephrotropic; antithyroid; dermatologic; immunomodulator;  
KW cell proliferation disorder; cell differentiation disorder;  
KW brain disorder; platelet disorder; breast disorder; colon disorder;  
KW kidney disorder; renal disorder; lung disorder; ovarian disorder;  
KW prostate disorder; cervical disorder; spleen disorder; thymus disorder;  
KW thyroid disorder; testes disorder; haematopoietic disorder;  
KW pancreatic disorder; skeletal muscle disorder; skin disorder;  
KW dermal disorder; bone metabolism disorder; immune disorder;  
KW inflammatory disorder; cardiovascular disorder;  
KW endothelial cell disorder; liver disorder; viral disease; pain disorder;  
KW metabolic disorder; neurological disorder;  
KW central nervous system disorder; erythroid disorder;  
KW blood vessel disorder; angiogenic disorder; cancer; heart failure;  
KW hypertension; angina; osteoarthritis; rheumatoid arthritis;  
KW multiple sclerosis; Crohn's disease; psoriasis; asthma;  
KW cell proliferation; cell differentiation; cell growth; cell division;  
KW human.

XX OS Homo sapiens.  
XX PN US2004058355-A1.  
XX PD 25-MAR-2004.  
XX PF 25-APR-2003; 2003US-00423543.  
XX 30-SEP-1998; 98US-00163821.  
PR 27-JAN-1999; 99US-0117580P.  
PR 25-MAR-1999; 99US-00276400.  
PR 30-JUL-1999; 99US-00365162.  
PR 09-SEP-1999; 99US-00392189.  
PR 05-OCT-1999; 99US-00412210.  
PR 23-NOV-1999; 99US-00448076.  
PR 29-FEB-2000; 2000US-0186061P.  
PR 28-APR-2000; 2000US-0200688P.  
PR 19-MAY-2000; 2000US-0205447P.  
PR 30-JUN-2000; 2000US-00608921.  
PR 31-JUL-2000; 2000US-0221925P.  
PR 25-SEP-2000; 2000US-0234922P.  
PR 25-SEP-2000; 2000US-0235035P.  
PR 08-NOV-2000; 2000US-0246669P.  
PR 09-NOV-2000; 2000US-00711216.  
PR 14-NOV-2000; 2000US-0248325P.  
PR 15-NOV-2000; 2000US-0248933P.  
PR 22-DEC-2000; 2000US-0257511P.  
PR 05-JAN-2001; 2001US-0260166P.  
PR 28-FEB-2001; 2001US-00797039.  
PR 27-APR-2001; 2001US-00845044.  
PR 20-JUL-2001; 2001US-00909743.  
PR 31-JUL-2001; 2001US-00920346.  
PR 13-AUG-2001; 2001US-00928531.  
PR 14-AUG-2001; 2001US-00929218.  
PR 15-AUG-2001; 2001US-0312539P.  
PR 25-SEP-2001; 2001US-00963159.  
PR 08-NOV-2001; 2001US-0008016.  
PR 13-NOV-2001; 2001US-00013055.  
PR 15-NOV-2001; 2001US-00003690.  
PR 30-JAN-2002; 2002US-00060763.  
PR 25-MAR-2002; 2002US-00105989.  
PR 12-APR-2002; 2002US-00121911.  
PR 12-AUG-2002; 2002US-00217168.  
PR 22-OCT-2002; 2002US-00278036.  
PR 02-JAN-2003; 2003US-00336489.  
PR 03-JAN-2003; 2003US-00336153.  
XX PA (MILL-) MILLENNIUM PHARM INC.  
XX PI Kapeller-Libermann R, Hunter JJ, Meyers RE, Rudolph-Owen LA;  
PI Curtis RAJ, Olandt PJ, Tsai F, Galvin KM, Chun M, Williamson MJ;  
PI Silos-Santiago I, Bandaru R;  
DR N-PSDB; ADL14160, ADL14162.  
XX WPI; 2004-268788/25.  
XX New 21910, 56634, 55053, 2504, 15977, 14760, 25501, 17903, 3700, 21529,  
PT 26176, 26343, 56638, 18610, 33217, 21967, h1983, m1983, 38555 or 593  
PT nucleic acid molecules and proteins, useful for treating, e.g. cancer,  
XX heart failure and angina.  
PS Claim 4; SEQ ID NO 44; 139pp; English.  
XX The invention describes an isolated 21910, 56634, 55053, 2504, 15977,  
CC 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217,  
CC 21967, h1983, m1983, 38555 or 593 nucleic acid molecule (I) comprising  
CC any one of 40 nucleotide sequences (I). The nucleic acid molecules and  
CC polypeptides are useful for diagnosing and treating a subject having a  
CC disorder, or a subject at risk of developing a disorder, which is  
CC associated with aberrant 21910, 56634, 55053, 2504, 15977, 14760, 25501,  
CC 17903, 3700, 21529, 26176, 26343, 56638, 18610, 33217, 21967, h1983,  
CC m1983, 38555 or 593 activity, such as cellular proliferative and/or

CC differentiative disorders, brain disorders, platelet disorders, breast  
CC disorders, colon disorders, kidney (renal) disorders, lung disorders,  
CC ovarian disorders, prostate disorders, cervical disorders, spleen  
CC disorders, thymus disorders, thyroid disorders, testes disorders,  
CC hematopoietic disorders, pancreatic disorders, skeletal muscle disorders,  
CC skin (dermal) disorders, disorders associated with bone metabolism,  
CC immune, e.g. inflammatory disorders, cardiovascular disorders,  
CC endothelial cell disorders, liver disorders, viral diseases, pain  
CC disorders, metabolic disorders, neurological or central nervous system  
CC disorders, erythroid disorders, blood vessel disorders or angiogenic  
CC disorders (all claimed), e.g. cancer, heart failure, hypertension,  
CC angina, osteoarthritis, rheumatoid arthritis, multiple sclerosis, Crohn's  
CC disease, psoriasis, or asthma. The nucleic acid molecules and  
CC polypeptides are also useful as modulating agents in regulating a variety  
CC of cellular processes, e.g. cell proliferation, differentiation, growth and  
CC division. This is the amino acid sequence of a novel human protein of the  
CC invention. Note: The sequences given in the specification are also  
CC available in electronic format from  
CC ftp.segdata.uspto.gov/sequence.html?docID=20040058355.  
XX  
SQ Sequence 628 AA;

Query Match 85.8%; Score 2824.5; DB 8; Length 628;  
Best Local Similarity 85.7%; Pred. No. 1.9e-239;  
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;  
QY 1 MESVALLQRPSPQSPASALASASARPLADGLIKSPKLMKKQAVKRHHKHLNLRHYEFL 60  
DB 1 MESLVFARRSGTPS-----AELARPLAELGLIKSPKLMKKQAVKRHHKHLNLRHYEFL 56  
QY 61 ETLGKGTGYGVKKARESSGRGLVAIKSRKDKIKDEQDLHRRREIEMSSLNPHIAT 120  
DB 57 ETLGKGTGYGVKKARESSGRGLVAIKSRKDKIKDEQDLHRRREIEMSSLNPHIAT 116  
QY 121 EVFENSSKIVIVMEYASRGDLYDIISERPLSRDARHFRQIVSAHYCHONGIVHRDL 180  
DB 117 EVFENSSKIVIVMEYASRGDLYDIISERQQLSREARHFRQIVSAHYCHONRVHRDL 176  
QY 181 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPVGVPEVDSWSL 240  
DB 177 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYTGVPEVDSWSL 236  
QY 241 GVLLYLIVHGTMPDGDHKLTVKQISNGAYREPPKPSDACGLIRWLLMVNPTRRATLED 300  
DB 237 GVLLYLIVHGTMPDGDHKLTVKQISNGAYREPPKPSDACGLIRWLLMVNPTRRATLED 296  
QY 301 VASHWWYNWGYTTGCGEQEALREGCHPSGDFGRASMAWDLRRSRPILLENKAVCSFFKQ 360  
DB 297 VASHWWYNWGYTRVGEQAPHEGGHFGSDSARASMAWDLRRSRPILLENKAVCSFFKQ 356  
QY 361 HVPGGGSTVPLERQHSKKSRKENDMAQNLQGDPAEDTSSRPGKSSLLKPLKGLKKSS 420  
DB 357 HAPGGGSTTGLERQHSKKSRKENDMAQSLHSDTADTAHRPGKSNLKLPGILKKVS 416  
QY 421 TSSEGEQEDPQELRPVPTDTPQVPVAVSLPLPRKGLKKSRQSGYSSPEPSESGELLD 480  
DB 417 ASAEQVEDPPELSPASPQAP---LLPKGILKKPRQSGYSSPEPSESGELLD 473  
QY 481 ASDVFVSGDPVEQKSPQASGLLHRRKILKNGKFSSTALETGTPFTGSLDQLASHPA 540  
DB 474 AGDVFFVSGDPKEQKPPQASGLLHRRKILKNGKFSSTALETGTPFTGSLDQLASHPA 533  
QY 541 ARSPRSGAVSDESILSSSEFDQLDLPERLPEPLRGCVSDNLRGLQEPSESG---LK 596  
DB 534 ARSPRSGAVSDESILSSSEFDQLDLPERLPEPLRGCVSDNLRGLQEPSESGSGCLR 593  
QY 597 RWMQESLGDSCFSLTDCQEVTAAYRQALGCSKLS 631  
DB 594 RWRQDPLGDSFSLTDCQEVTAAYRQALVCSKLT 628

RESULT 11  
ADO20172

ID ADO20172 standard; protein; 628 AA.  
XX ADO20172;  
AC  
XX  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human PRO polypeptide #540.  
XX  
XX Human; PRO; immune related disorder; systemic lupus erythematosus;  
KW rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;  
KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;  
KW autoimmune hemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;  
KW diabetes mellitus; renal disease; demyelinating disease;  
KW central nervous system; peripheral nervous system;  
KW demyelinating polyneuropathy; Guillain-Barre syndrome;  
KW chronic inflammatory demyelinating polyneuropathy.  
XX  
OS Homo sapiens.  
XX  
XX WO2004043361-A2.  
XX  
XX 27-MAY-2004.  
XX  
XX 06-NOV-2003; 2003WO-US035268.  
XX  
XX 08-NOV-2002; 2002US-0425235P.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;  
PI Wood WI, Wu TD;  
XX  
XX WPI; 2004-420067/39.  
XX  
XX N-PSDB; ADO20171.  
XX  
XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for  
PT treating an immune related disorder such as systemic lupus erythematosus,  
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or  
PT spondyloarthropathy.  
XX  
XX  
PS Claim 7; SEQ ID NO 1080; 1731pp; English.  
XX  
XX The invention relates to human PRO polypeptides and the polynucleotides  
CC encoding them. The polypeptides and polynucleotides are useful for  
CC treating and diagnosing immune related disorders in mammals. The immune  
CC related disorders include systemic lupus erythematosus, rheumatoid  
CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic  
CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune  
CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes  
CC mellitus, immune-mediated renal disease, demyelinating diseases of the  
CC central or peripheral nervous system, demyelinating polyneuropathy,  
CC Guillain-Barre syndrome and chronic inflammatory demyelinating  
CC polyneuropathy. This sequence represents a human PRO polypeptide of the  
CC invention.  
XX  
SQ Sequence 628 AA;

Query Match 85.8%; Score 2824.5; DB 8; Length 628;  
Best Local Similarity 85.7%; Pred. No. 1.9e-239;  
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;  
QY 1 MESVALLQRPSPQSPASALASASARPLADGLIKSPKLMKKQAVKRHHKHLNLRHYEFL 60  
DB 1 MESLVFARRSGTPS-----AELARPLAELGLIKSPKLMKKQAVKRHHKHLNLRHYEFL 56  
QY 61 ETLGKGTGYGVKKARESSGRGLVAIKSRKDKIKDEQDLHRRREIEMSSLNPHIAT 120  
DB 57 ETLGKGTGYGVKKARESSGRGLVAIKSRKDKIKDEQDLHRRREIEMSSLNPHIAT 116  
QY 121 EVFENSSKIVIVMEYASRGDLYDIISERPLSRDARHFRQIVSAHYCHONGIVHRDL 180  
DB 117 EVFENSSKIVIVMEYASRGDLYDIISERQQLSREARHFRQIVSAHYCHONRVHRDL 176

QY 181 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYGPEVDSWSL 240  
DB 177 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYGPEVDSWSL 236  
QY 241 GVLLYILVHGTMPDGDHKLTVKQISNGAYREPPKSDACGLIRWLLMVNPTTRATLED 300  
DB 237 GVLLYILVHGTMPDGDHKLTVKQISNGAYREPPKSDACGLIRWLLMVNPTTRATLED 296  
QY 301 VASHWVNWGYTTGCGEALREGGHPGSDGFRASMDLRRSRPILLENKAKVCSFFKQ 360  
DB 297 VASHWVNWGYTTGCGEALREGGHPGSDGFRASMDLRRSRPILLENKAKVCSFFKQ 356  
QY 361 HVPGGGTVPLERQHSILKSKRKNDAQNLQDPAEDTSSRPGKSSILKPKGILKKSS 420  
DB 357 HAPGGGTTPLERQHSILKSKRKNDAQNLQDPAEDTSSRPGKSSILKPKGILKKSS 416  
QY 421 TSSEGEQDPELRPVDPDTPGQPVAVSLLPRKGLKKRQRESGYSSPEPSSGELL 480  
DB 417 ASAEVQEDPELSPIPASPGQAP---LLPKGILKKRQRESGYSSPEPSSGELL 473  
QY 481 ASDVFGSDPVEOKSPQASGLLHHRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540  
DB 474 AGDVFGSDPVEOKSPQASGLLHHRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 533  
QY 541 ARPSRPSGAVSEDSILSESFDQLDLPERLPETPLRGCVSDNLRGLEQPPSEG---LK 596  
DB 534 ARASRPSGAVSEDSILSESFDQLDLPERLPETPLRGCVSDNLRGLEQPPSEG---LK 593  
QY 597 RWOESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631  
DB 594 RWRQDPLGDSFSLTDCQEVTAAYRQALGICSKLT 628

## RESULT 12

ADJ96620  
ID ADJ96620 standard; protein; 672 AA.  
XX  
AC ADJ96620;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Human calcium/calmodulin-dependent protein kinase NuaK2 protein SeqID 77.  
XX  
KW kinase; human; tyrosine protein kinase; serine/threonine protein kinase;  
KW PK; STK; gene therapy; cancer; immune-related disease;  
KW cardiovascular disease; brain; neuronal associated disease; metabolic;  
KW inflammatory disorder; cytosolic; neuroprotective; immunomodulator;  
KW antiinflammatory; enzyme; calcium/calmodulin-dependent protein kinase;  
KW NuaK2.  
XX  
OS Homo sapiens.  
OS 72.  
XX  
PN WO2004006838-A2.  
XX  
PD 22-JAN-2004.  
XX  
PF 15-JUL-2003; 2003WO-US021730.  
XX  
PR 15-JUL-2002; 2002US-0395632P.  
XX  
PA (SUGEN-) SUGEN INC.  
XX  
PI Whyte D, Manning G, Caenepeel S;  
XX  
DR WPI; 2004-122753/12.  
DR N-PSDB; ADJ96554.  
XX  
PT New nucleic acid molecule encoding a kinase polypeptide, useful for  
PT preparing a composition for treating diseases or disorders, e.g., cancer,  
PT or neurological, immunological or inflammatory disorders.  
XX  
PS Claim 1; SEQ ID NO 77; 366pp; English.

XX This invention relates to a novel isolated, enriched or purified nucleic  
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates  
CC to human tyrosine and serine/threonine protein kinases (PTK's and STK's),  
CC as well as protein kinase-like enzymes. The present invention describes  
CC screening methods to identify agonists, antagonists and antibodies that  
CC can be used to modulate the activity or function of the mammalian kinase  
CC enzymes. As such, these compositions can be used for gene therapy  
CC purposes to treat diseases or disorders including cancer, immune-related  
CC diseases, cardiovascular disease, brain or neuronal associated disease,  
CC metabolic and inflammatory disorders. Accordingly, they exhibit  
CC cytosolic, neuroprotective, immunomodulator and antiinflammatory  
CC activities. This polypeptide sequence is a human kinase protein sequence  
CC of the invention.

SQ Sequence 672 AA;

Query Match 85.8%; Score 2824.5; DB 8; Length 672;  
Best Local Similarity 85.7%; Pred. No. 2.1e-239;  
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;

QY 1 MESVALLQRFSPQAPSASALASERPLADGLIKSPKPLMKKQAVKRHHKHLRHYEFL 60  
DB 45 MESLVFARRSGTPTS---AAELARPLAELGLIKSPKPLMKKQAVKRHHKHLRHYEFL 100  
QY 61 ETGLKGYGVKKARESSGRVLAIKIRKDKIKDEQDOLLHRRREIEMSSLNHPHIAIH 120  
DB 101 ETGLKGYGVKKARESSGRVLAIKIRKDKIKDEQDOLLHRRREIEMSSLNHPHIAIH 160  
QY 121 EVFENSISKIVIMVMEYASRGDLYDIYSERPLSRDARHFRQIVSALHYCHQNGIVHRDL 180  
DB 161 EVFENSISKIVIMVMEYASRGDLYDIYSERQQLSREARHFRQIVSAVHYCHQNRVHRDL 220  
QY 181 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYGPEVDSWSL 240  
DB 221 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYGPEVDSWSL 280  
QY 241 GVLLYILVHGTMPDGDHKLTVKQISNGAYREPPKSDACGLIRWLLMVNPTTRATLED 300  
DB 281 GVLLYILVHGTMPDGDHKLTVKQISNGAYREPPKSDACGLIRWLLMVNPTTRATLED 340  
QY 301 VASHWVNWGYTTGCGEALREGGHPGSDGFRASMDLRRSRPILLENKAKVCSFFKQ 360  
DB 341 VASHWVNWGYTTGCGEALREGGHPGSDGFRASMDLRRSRPILLENKAKVCSFFKQ 400  
QY 361 HVPGGGTVPLERQHSILKSKRKNDAQNLQDPAEDTSSRPGKSSILKPKGILKKSS 420  
DB 401 HAPGGGTTPLERQHSILKSKRKNDAQNLQDPAEDTSSRPGKSSILKPKGILKKSS 460  
QY 421 TSSEGEQDPELRPVDPDTPGQPVAVSLLPRKGLKKRQRESGYSSPEPSSGELL 480  
DB 461 ASAEVQEDPELSPIPASPGQAP---LLPKGILKKRQRESGYSSPEPSSGELL 517  
QY 481 ASDVFGSDPVEOKSPQASGLLHHRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540  
DB 518 AGDVFGSDPVEOKSPQASGLLHHRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 577  
QY 541 ARPSRPSGAVSEDSILSESFDQLDLPERLPETPLRGCVSDNLRGLEQPPSEG---LK 596  
DB 578 ARASRPSGAVSEDSILSESFDQLDLPERLPETPLRGCVSDNLRGLEQPPSEG---LK 637  
QY 597 RWOESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631  
DB 638 RWRQDPLGDSFSLTDCQEVTAAYRQALGICSKLT 672

## RESULT 13

AAB65632  
ID AAB65632 standard; protein; 629 AA.  
XX  
AC AAB65632;  
XX  
DT 27-MAR-2001 (first entry)







CC of a test compound and in gene therapy. The present sequence is human  
CC PKIN-12 protein

XX  
SQ Sequence 594 AA;

```
Query Match      82.6%; Score 2720.5; DB 5; Length 594;
Best Local Similarity 87.3%; Pred. No. 2.6e-230;
Matches 521; Conservative 22; Mismatches 47; Indels 7; Gaps 2;

Qy 39 MKQAVRHHHKNLRRHYELETIGKTYGKVKKARESSGRLVAIKSIRKDKIKDEODL 98
Db 1 MKQAVRHHHKNLRRHYELETIGKTYGKVKKARESSGRLVAIKSIRKDKIKDEODL 60

Qy 99 LHIREIEIMSSLNHPHIIAIHEVFENSKKIVIMEYASRGDLVDYISERPRLSERDARH 158
Db 61 MHIREIEIMSSLNHPHIIAIHEVFENSKKIVIMEYASRGDLVDYISERQQLSEREARH 120

Qy 159 FFRQIVSALHYCHONGIVHRDLKLENILLDANGNIKIADFGLSNLYHKGFQTFCSPL 218
Db 121 FFRQIVSAYHYCHQNRVVHRDLKLENILLGANGNIKIADFGLSNLYHQGFQTFCSPL 180

Qy 219 YASPEIVNGKPYGVPEVDSWSLGLVLLYLVHGTWPFQDQHKTLVKQISNGAYREPPKPS 278
Db 181 YASPEIVNGKPYGVPEVDSWSLGLVLLYLVHGTWPFQDQHKTLVKQISNGAYREPPKPS 240

Qy 279 DACGLIRLLMVNPTRRATLEDAVSHWVNWNGYTTGVGEQALREGGHPGSGDFGRASMD 338
Db 241 DACGLIRLLMVNPTRRATLEDAVSHWVNWNGYATRVGEQAPHEGGHPGSGDSARASMD 300

Qy 339 WLRSSRPILLENGAKVCSFFKQHVPGGGSTVPGLERQHSLSKSRKENDMAQNLOGDPAED 398
Db 301 WLRSSRPILLENGAKVCSFFKQHAPGGSTTPGLERQHSLSKSRKENDMAQSLHSDTADD 360

Qy 399 TSSRPKSSLLPKGILKKSSSTSSGEVQEDPQELRPVDPDPGPVPAVSLIPKGIKK 458
Db 361 TAHRPGKSNLKLPGILKKVSASAEVQEDPPELSPIPASPGQAP---LLPKKGIKK 417

Qy 459 SQRSEGYYSPEPSESGELLDSDFVSGDPVQKSPQASGLLHKKGIKLGKFTSRT 518
Db 418 PRQSEGYYSPEPSESGELLDSDFVSGDPKPKPPQASGLLHKKGIKLGKFSQT 477

Qy 519 ALEGTTPSTFGSLDQLASSHPAARPSRGAVSEDSILSSSEFPQDLPERLPETPLRGC 578
Db 478 ALELAAPTTFGSLDELAPRPLARASRPSGAVSEDSILSSSEFPQDLPERLPETPLRGC 537

Qy 579 VSDNLRGLQPPSEG---LKRWQESLGDSCFSLTDCQEVTAAYQALGICSKLS 631
Db 538 VSDNLTGLEPPEPSEPGSCLRRWRQDPLGDSFSLTDCQEVTAAYQALRVCSKLT 594
```

Search completed: January 22, 2005, 06:38:26  
Job time : 165.178 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 06:30:32 ; Search time 40.6867 Seconds  
(without alignments)  
1028.511 Million cell updates/sec

Title: US-09-980-464-11  
Perfect score: 3293  
Sequence: 1 MESVALLQRPSPASASALA.....DCQEVTAAYRQALGICKLS 631

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3293	100.0	631	4	US-09-579-664B-11
2	3293	100.0	631	4	US-10-355-975A-11
3	3278.5	99.6	630	4	US-10-355-975A-38
4	703.5	21.4	729	2	US-08-677-298-2
5	703.5	21.4	729	4	US-09-523-849-33
6	694.5	21.1	713	4	US-09-538-092-1022
7	691.5	21.0	793	4	US-09-523-849-32
8	688	20.9	745	4	US-09-523-849-36
9	680	20.6	724	4	US-09-984-890-2
10	680	20.6	724	4	US-10-274-194-2
11	677.5	20.6	722	4	US-09-984-890-4
12	677.5	20.6	722	4	US-10-274-194-4
13	675	20.5	779	4	US-08-817-832B-31
14	673.5	20.5	722	4	US-08-817-832B-32
15	653	19.8	604	4	US-09-523-849-35
16	648.5	19.7	633	3	US-08-557-006C-43
17	648.5	19.7	633	4	US-09-538-092-212
18	648.5	19.7	633	4	US-09-633-328B-3
19	648.5	19.7	633	4	US-09-824-735-3
20	648.5	19.7	1203	4	US-09-799-875-5
21	645.5	19.6	512	4	US-09-633-328B-2
22	636.5	19.3	552	4	US-09-824-735-4
23	635.5	19.3	345	3	US-09-101-146-1
24	635.5	19.3	345	3	US-09-523-849-34
25	631	19.2	511	4	US-09-633-328B-4
26	628.5	19.1	552	4	US-09-538-092-1212
27	625.5	19.0	418	4	US-09-248-796A-18441

28	625	19.0	257	3	US-09-101-146-6	Sequence 6, Appli
29	623.5	18.9	552	3	US-08-557-006C-40	Sequence 40, Appl
30	619.5	18.8	257	2	US-07-857-224B-25	Sequence 25, Appl
31	613.5	18.6	778	4	US-10-116-326-2	Sequence 2, Appli
32	613.5	18.6	778	4	US-10-003-690-2	Sequence 2, Appli
33	599.5	18.2	668	4	US-09-930-181-2	Sequence 2, Appli
34	549	16.7	1064	4	US-09-538-092-154	Sequence 154, App
35	547	16.6	703	4	US-10-116-326-6	Sequence 6, Appli
36	528.5	16.0	603	4	US-09-930-181-17	Sequence 17, Appl
37	525	15.9	260	2	US-07-857-224B-28	Sequence 28, Appl
38	522	15.9	776	1	US-08-198-446B-17	Sequence 17, Appl
39	522	15.9	776	2	US-08-870-693-17	Sequence 17, Appl
40	522	15.9	821	1	US-08-198-446B-6	Sequence 6, Appli
41	522	15.9	821	2	US-08-870-693-6	Sequence 6, Appli
42	519	15.8	391	4	US-09-248-796A-18435	Sequence 18435, A
43	515.5	15.7	160	4	US-09-270-767-32886	Sequence 32886, A
44	514	15.6	915	4	US-09-538-092-63	Sequence 63, Appl
45	513	15.6	260	2	US-07-857-224B-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1  
US-09-579-664B-11  
; Sequence 11, Application US/09579664B  
; Patent No. 6514719  
; GENERAL INFORMATION:  
; APPLICANT: Immunex Corporation  
; APPLICANT: Bird, Timothy A.  
; APPLICANT: Virca, G. Duke  
; APPLICANT: Martin, Unja  
; APPLICANT: Anderson, Dirk M.  
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES  
; FILE REFERENCE: 2923-A  
; CURRENT APPLICATION NUMBER: US/09/579,664B  
; CURRENT FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 11  
; LENGTH: 631  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; US-09-579-664B-11

Query Match	100.0%;	Score	3293;	DB	4;	Length	631;
Best Local Similarity	100.0%;	Pred. NO.	2.2e-277;	Mismatches	0;	Indels	0;
Matches	631;	Conservative	0;	0;	0;	Gaps	0;
Qy	1	MESVALLQRPSPASASALASARPLADGLIKSPKPLMKQAVKRHHKHLRHYEFL	60				
Db	1	MESVALLQRPSPASASALASARPLADGLIKSPKPLMKQAVKRHHKHLRHYEFL	60				
Qy	61	ETLGKGYGVKKARSSGRBLVAIKSRKDKIKDEQLLHRRREIEMSSLNHPHIIAIH	120				
Db	61	ETLGKGYGVKKARSSGRBLVAIKSRKDKIKDEQLLHRRREIEMSSLNHPHIIAIH	120				
Qy	121	EVFENSSKIVIVMEYASRGDLYISERPLSERDARHFRQIVSALHYCHONGIVHRDL	180				
Db	121	EVFENSSKIVIVMEYASRGDLYISERPLSERDARHFRQIVSALHYCHONGIVHRDL	180				
Qy	181	KLENILLDANGNKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPVDSWSL	240				
Db	181	KLENILLDANGNKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPVDSWSL	240				
Qy	241	GVLlyLVHGTMPFDGQDHKTLVKQISNGAYRPPKPSDACGLIRWLLMVNPTRRATLED	300				
Db	241	GVLlyLVHGTMPFDGQDHKTLVKQISNGAYRPPKPSDACGLIRWLLMVNPTRRATLED	300				
Qy	301	VASHWWNWGYYTTCVGEQEARREGCHPSGDFGRASMDWLRSSRPILLECAKVCSPFKQ	360				
Db	301	VASHWWNWGYYTTCVGEQEARREGCHPSGDFGRASMDWLRSSRPILLECAKVCSPFKQ	360				

```
Qy 361 HVPGGGTVPLGRQHSLLKSRKENDMAQNLOQDPAEDTSRPGKSSLLPKGILKKSS 420
Db 361 HVPGGGTVPLGRQHSLLKSRKENDMAQNLOQDPAEDTSRPGKSSLLPKGILKKSS 420
Qy 421 TSSGEVQEDPQELRPVPTPGQVPVAVSLLPRKGILKKSRQREGSYSSPPSPSGELL 480
Db 421 TSSGEVQEDPQELRPVPTPGQVPVAVSLLPRKGILKKSRQREGSYSSPPSPSGELL 480
Qy 481 ASDVVFVSGDPVEQKSPQASGLLHRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540
Db 481 ASDVVFVSGDPVEQKSPQASGLLHRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540
Qy 541 ARPSRPSGAVSEDSILSSESDQLDLPERLPETPLRGCVSDNLRGLQEPSEGLKRWQ 600
Db 541 ARPSRPSGAVSEDSILSSESDQLDLPERLPETPLRGCVSDNLRGLQEPSEGLKRWQ 600
Qy 601 ESLGDSFSLTDCQEVTAAYROALGICSKLS 631
Db 601 ESLGDSFSLTDCQEVTAAYROALGICSKLS 631

RESULT 2
US-10-355-975A-11
; Sequence 11, Application US/10355975A
; Patent No. 6759223
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: CALCIUM/CALMODULIN-DEPENDENT KINASE
; FILE REFERENCE: 2923-B
; CURRENT APPLICATION NUMBER: US/10/355.975A
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Mus musculus
; ORGANISM: Mus musculus
US-10-355-975A-11

Query Match 100.0%; Score 3293; DB 4; Length 631;
Best Local Similarity 100.0%; Pred. No. 2.2e-277;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESVALLQRPSPQAPSASALASESARPLADGLIKSPKPLMKKQAVKRHHKHNLRHYEFL 60
Db 1 MESVALLQRPSPQAPSASALASESARPLADGLIKSPKPLMKKQAVKRHHKHNLRHYEFL 60
Qy 61 ETLGKGTGKVKKARESGRLVAIKSRDKIKDEQDLHRRREIETMSSLNHPHIIAIIH 120
Db 61 ETLGKGTGKVKKARESGRLVAIKSRDKIKDEQDLHRRREIETMSSLNHPHIIAIIH 120
Qy 121 EVFENSISKIVMEYASRGDLYDIYSERPLSRDARHFFRQIVSALHYCHQNGIVHRDL 180
Db 121 EVFENSISKIVMEYASRGDLYDIYSERPLSRDARHFFRQIVSALHYCHQNGIVHRDL 180
Qy 181 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPVEDWSL 240
Db 181 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPVEDWSL 240
Qy 241 GVLLYILVHGTMPDQDHKTLVKQISNGAYREPPEKPSDACGLIRLWMNPTTRATLED 300
Db 241 GVLLYILVHGTMPDQDHKTLVKQISNGAYREPPEKPSDACGLIRLWMNPTTRATLED 300
Qy 301 VASHWWNVGTYTTCVGQEQALREGGHPSGDGRASMDLRRSRPLLENGAKVCSFFKQ 360
Db 301 VASHWWNVGTYTTCVGQEQALREGGHPSGDGRASMDLRRSRPLLENGAKVCSFFKQ 360
Qy 361 HVPGGGTVPLGRQHSLLKSRKENDMAQNLOQDPAEDTSRPGKSSLLPKGILKKSS 420
Db 361 HVPGGGTVPLGRQHSLLKSRKENDMAQNLOQDPAEDTSRPGKSSLLPKGILKKSS 420
```

```
Db 361 HVPGGGTVPLGRQHSLLKSRKENDMAQNLOQDPAEDTSRPGKSSLLPKGILKKSS 420
Qy 421 TSSGEVQEDPQELRPVPTPGQVPVAVSLLPRKGILKKSRQREGSYSSPPSPSGELL 480
Db 421 TSSGEVQEDPQELRPVPTPGQVPVAVSLLPRKGILKKSRQREGSYSSPPSPSGELL 480
Qy 481 ASDVVFVSGDPVEQKSPQASGLLHRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540
Db 481 ASDVVFVSGDPVEQKSPQASGLLHRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540
Qy 541 ARPSRPSGAVSEDSILSSESDQLDLPERLPETPLRGCVSDNLRGLQEPSEGLKRWQ 600
Db 541 ARPSRPSGAVSEDSILSSESDQLDLPERLPETPLRGCVSDNLRGLQEPSEGLKRWQ 600
Qy 601 ESLGDSFSLTDCQEVTAAYROALGICSKLS 631
Db 601 ESLGDSFSLTDCQEVTAAYROALGICSKLS 631

RESULT 3
US-10-355-975A-38
; Sequence 38, Application US/10355975A
; Patent No. 6759223
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: CALCIUM/CALMODULIN-DEPENDENT KINASE
; FILE REFERENCE: 2923-B
; CURRENT APPLICATION NUMBER: US/10/355.975A
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Mus musculus
; ORGANISM: Mus musculus
US-10-355-975A-38

Query Match 99.6%; Score 3278.5; DB 4; Length 630;
Best Local Similarity 99.8%; Pred. No. 4.1e-276;
Matches 630; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 MESVALLQRPSPQAPSASALASESARPLADGLIKSPKPLMKKQAVKRHHKHNLRHYEFL 60
Db 1 MESVALLQRPSPQAPSASALASESARPLADGLIKSPKPLMKKQAVKRHHKHNLRHYEFL 60
Qy 61 ETLGKGTGKVKKARESGRLVAIKSRDKIKDEQDLHRRREIETMSSLNHPHIIAIIH 120
Db 61 ETLGKGTGKVKKARESGRLVAIKSRDKIKDEQDLHRRREIETMSSLNHPHIIAIIH 120
Qy 121 EVFENSISKIVMEYASRGDLYDIYSERPLSRDARHFFRQIVSALHYCHQNGIVHRDL 180
Db 121 EVFENSISKIVMEYASRGDLYDIYSERPLSRDARHFFRQIVSALHYCHQNGIVHRDL 180
Qy 181 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPVEDWSL 240
Db 181 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPVEDWSL 240
Qy 241 GVLLYILVHGTMPDQDHKTLVKQISNGAYREPPEKPSDACGLIRLWMNPTTRATLED 300
Db 241 GVLLYILVHGTMPDQDHKTLVKQISNGAYREPPEKPSDACGLIRLWMNPTTRATLED 300
Qy 301 VASHWWNVGTYTTCVGQEQALREGGHPSGDGRASMDLRRSRPLLENGAKVCSFFKQ 360
Db 301 VASHWWNVGTYTTCVGQEQALREGGHPSGDGRASMDLRRSRPLLENGAKVCSFFKQ 360
Qy 361 HVPGGGTVPLGRQHSLLKSRKENDMAQNLOQDPAEDTSRPGKSSLLPKGILKKSS 420
Db 361 HVPGGGTVPLGRQHSLLKSRKENDMAQNLOQDPAEDTSRPGKSSLLPKGILKKSS 420
```

Qy 421 TSSEVEDQDQELRPVDTTGPQVPVAVSLPRKGIKKSRQREGSYSSPEPSGELL 480  
Db 420 TSSEVEDQDQELRPVDTTGPQVPVAVSLPRKGIKKSRQREGSYSSPEPSGELL 479  
Qy 481 ASDVFVSGDPVEQKSPQASGLLHHRKGIKKLNGKFSRTALEGTTPTFGSLDQLASSHPA 540  
Db 480 ASDVFVSGDPVEQKSPQASGLLHHRKGIKKLNGKFSRTALEGTTPTFGSLDQLASSHPA 539  
Qy 541 ARPSRPGAVSEDSILSESFQDLQDLPERLPETPLRGCVSDNLRGLQEPSEGLKRWQ 600  
Db 540 ARPSRPGAVSEDSILSESFQDLQDLPERLPETPLRGCVSDNLRGLQEPSEGLKRWQ 599  
Qy 601 ESLGDSFSLTDCQEVTAAYRQALGICSKLS 631  
Db 600 ESLGDSFSLTDCQEVTAAYRQALGICSKLS 630

## RESULT 4

US-08-677-298-2  
; Sequence 2, Application US/08677298  
; Patent No. 5863729  
; GENERAL INFORMATION:  
; APPLICANT: Pivnicka-Worms, Helen  
; TITLE OF INVENTION: DNA SEQUENCES ENCODING HUMAN TCAK-1  
; TITLE OF INVENTION: KINASE  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle, Suite 201  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: USA  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/677,298  
; FILING DATE: 09-JUL-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Caruthers, Jennie M.  
; REGISTRATION NUMBER: 34,464  
; REFERENCE/DOCKET NUMBER: 9-96  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 729 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

## US-08-677-298-2

Query Match 21.4%; Score 703.5; DB 2; Length 729;  
Best Local Similarity 35.1%; Pred. No. 3.4e-52;  
Matches 195; Conservative 76; Mismatches 180; Indels 105; Gaps 18;  
Qy 57 YFELETIGKGTGKVKKARE--SSGRLVAIKSIRKDKIKDEQDLHRRREIEMSLNHPH 115  
Db 56 YRLKTKTGKGNFAKVLARHILTGREVAIKIDKTLQ--NPTSLOKLFRVIRIMKILNHPN 114  
Qy 116 IIAIHEVFENSISKIVIMVEASRGDLYDIYSERPLSERDARHPFRQIVSALHYCHQNGI 175  
Db 115 IVKLFEVETIEKTLVLYMEYASGGEVFDYLVHGRMKEKARSFRQIVSAVQYCHQKRI 174  
Qy 176 VHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASVEIVNGKPYGVPEV 235  
Db 175 VHRDLKAENLLDADNMNIKIADFGFSNEFTVGGKLDTCGSPPPYAAPELFGKKYDGPV 234

Qy 236 DWSLGLVLLYILVHGTMPFPDQDHKTLVKQISNGAYREPCKPSDAC--GLIRWLLMVNPT 294  
Db 235 DVNSLGVILYTLVSGSLPFDQNLKELRERVLGKRYIPFVMSDTCENLLKRFVLNPIK 294  
Qy 295 RATLEDVASHWVNWGVTGVEQEAALREGHSGDPRGRASMDWLRSSRPILLENGAKV 354  
Db 295 RGLEQIMKORWINAGH-----BEDELKPFVPELIDISDQKRIDIM-----V 336  
Qy 355 CSFFKQHVPGGSGTVPCGLERQHSIKSRKENDMAQNL---QGDPAEDTSRPGKSSLKLP 411  
Db 337 GMYSQB-----EIQESLSKMKYDEITATYLLLRKSELDASSSSSSLSLA 385  
Qy 412 KGILKKKSTSSGE-----VQ---EDQELRPVDPDTPGQVPVAVSLPRK-----GIL 456  
Db 386 KVRPSSDLNNTGQSPHHKQVRSVSSSQKORRYSDHAGPAIPSVVAYPKRSQTSABDGL 445  
Qy 457 KK---SQRES-----GYSSPEPESGEGELLDASDVFVSGDPVEQKSPQ 497  
Db 446 KEDGISRSKSSGSAVGKGIAPASPMLGASNPKNKADIPERKKSSTV-----PSSNT 497  
Qy 498 ASGLLLHRK-----GILKNGKFSRTALEGTTPTFGSLDQLASSH---PAA 541  
Db 498 ASGMRTRTYVCERTTADRHVSIQ--NGKENSTIPDQRTP-----VASTHSISSAA 548  
Qy 542 RPSR---PSGAVSEDS 554  
Db 549 TPDRIREFRGTASRST 564

## RESULT 5

US-09-523-849-33  
; Sequence 33, Application US/09523849  
; Patent No. 6458561  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Molteni, Angela  
; APPLICANT: Magnaghi, Paola  
; APPLICANT: Bosotti, Roberta  
; APPLICANT: Scaccheri, Emanuela  
; APPLICANT: Isacchi, Antonella  
; APPLICANT: Hodgson, Dave  
; TITLE OF INVENTION: HUMAN NIM1 KINASE  
; FILE REFERENCE: PC-0009 US  
; CURRENT APPLICATION NUMBER: US/09/523,849  
; CURRENT FILING DATE: 2000-03-13  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PERL Program  
; SEQ ID NO 33  
; LENGTH: 729  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: GenBank Accession No. 6458561 g3089349  
US-09-523-849-33

Query Match 21.4%; Score 703.5; DB 4; Length 729;  
Best Local Similarity 35.1%; Pred. No. 3.4e-52;  
Matches 195; Conservative 76; Mismatches 180; Indels 105; Gaps 18;  
Qy 57 YFELETIGKGTGKVKKARE--SSGRLVAIKSIRKDKIKDEQDLHRRREIEMSLNHPH 115  
Db 56 YRLKTKTGKGNFAKVLARHILTGREVAIKIDKTLQ--NPTSLOKLFRVIRIMKILNHPN 114  
Qy 116 IIAIHEVFENSISKIVIMVEASRGDLYDIYSERPLSERDARHPFRQIVSALHYCHQNGI 175  
Db 115 IVKLFEVETIEKTLVLYMEYASGGEVFDYLVHGRMKEKARSFRQIVSAVQYCHQKRI 174  
Qy 176 VHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASVEIVNGKPYGVPEV 235  
Db 175 VHRDLKAENLLDADNMNIKIADFGFSNEFTVGGKLDTCGSPPPYAAPELFGKKYDGPV 234  
Qy 236 DWSLGLVLLYILVHGTMPFPDQDHKTLVKQISNGAYREPCKPSDAC--GLIRWLLMVNPT 294

Db 235 DVMSGLVILYTLVSGSLPFDGQNLKELRVRVLRKYRIPFYMSTDCENLLKRFVLVLPK 294  
Qy 295 RATLEDVASHWVNWGTYTGVGEQALREGCHPSGDFGRASMDWLRSSRPILLECAKV 354  
Db 295 RGTLEQIMKDRWINAGH-----EDELKPFVEPELDISDQKRIDIM-----V 336  
Qy 355 CSFPKOHVPGGGSTVPGLEQHSLLKSKRKNDAQN-----QGDPAEDTSSRPGKSSIKLP 411  
Db 337 GMGYSQE-----EIQESLSKWKYDEITATYLLLRKSKSEVPPSSDLNNSQS- 384  
Qy 412 KGLKXKSSSSGE-----VQ-----EDPQLRPVDPDTPQPVPAVSLLPK-----GIL 456  
Db 386 KVRPSSDLNNSGSPHHKQVRSVSSQKQRRYSDHAGPALPSVAVPKRSQTSADGDL 445  
Qy 457 KK-----SRQRES-----GYVSSPSPSESGELLDASDVFSVGDVPQKSPQ 497  
Db 446 KEDGISRSKSSGSAVGGKIAPASPM--LGNASPNKADIPERKKSSTV-----PSSNT 497  
Qy 498 ASGULLHRK-----GILKNGKFSRTALEGTTPTSTFGSLDQLASH-----PAA 541  
Db 498 ASGWMTRNTYVCSERTTDRHSVIQ-NGKENSTIPDQRT-----VASTHSISSAA 548  
Qy 542 RPSR---PSGAVSEDS 554  
Db 549 TPDRIFRGTASRST 564

## RESULT 6

US-09-538-092-1022  
; Sequence 1022, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Mansfield, Traci A.  
; APPLICANT: Giot, Loic  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538.092  
; CURRENT FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CuratSeqFormatter Version 0.9  
; SEQ ID NO 1022  
; LENGTH: 713  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: Polypeptide Accession Number P27448  
US-09-538-092-1022

Query Match 21.1%; Score 694.5; DB 4; Length 713;  
Best Local Similarity 34.5%; Pred. No. 2e-51;  
Matches 191; Conservative 75; Mismatches 171; Indels 117; Gaps 17;  
Qy 57 YEFLETIGKTYGKVKARE--SSGRLVAIKSRDKIKDEQDLHRRREIEMSLNHPH 115  
Db 56 YRLKTTKGNFAVKLARHLTLTGREVAIKIDKTL-NPTSLOKLFREVRIMKILNHPN 114  
Qy 116 IIAIHEVFENSCKTIVMEYASRGDLYDIYSERPLSERDARHFFRQIVSALHYCHONGI 175  
Db 115 IVKLFEVETIOKTLIMEYASGGKVFDFYLAHGRMKKEARSFRQIVSAVQYCHOKRI 174  
Qy 176 VHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYGVPEV 235  
Db 175 VHRDLKAENLLDADNMKIADFGFSNEFTVGGKLDTPFCGSPPYAAPELFGKYYDGPV 234  
Qy 236 DSWSLGVLLYILVHGTWMPFDGQDHKTLVKQISNGAYREPPEKPSDAC-GLIRWLLMVNPT 294

Db 235 DVMSGLVILYTLVSGSLPFDGQNLKELRVRVLRKYRIPFYMSTDCENLLKRFVLVLPK 294  
Qy 295 RATLEDVASHWVNWGTYTGVGEQALREGCHPSGDFGRASMDWLRSSRPILLECAKV 354  
Db 295 RGTLEQIMKDRWINAGH-----EDELKPFVEPELDISDQKRIDIM-----V 336  
Qy 355 CSFPKOHVPGGGSTVPGLEQHSLLKSKRKNDAQN-----QGDPAEDTSSRPGKSS 407  
Db 337 GMGYSQE-----EIQESLSKWKYDEITATYLLLRKSKSEVPPSSDLNNSQS- 384  
Qy 408 LKLPKGLKXKSSSTSGEVEDPQELRPVDPDTPQPVPAVSLLPK----- 452  
Db 385 ---PHHKVQSVSS-----QKQRRYSDHAGPISVAVPKRSQTSADSLKEDG 433  
Qy 453 -----KGLKXKQRESGYVSSPSPSESGELLDASDVFSVGDVPQKSPQAS 499  
Db 434 ISSRKSTGSAVGGKIAPASPM--LGNASPNKADIPERKKSSTV-----PSSNTAS 483  
Qy 500 GLLLRK-----GILKNGKFSRTALEGTTPTSTFGSLDQLASH-----PAARP 543  
Db 484 GWMTRNTYVCSERTTDRHSVIQ-NGKENSTIPDQRT-----VASTHSISSAATP 534  
Qy 544 SR---PSGAVSEDS 554  
Db 535 DRIRFRGTASRST 548

## RESULT 7

US-09-523-849-32  
; Sequence 32, Application US/09523849  
; Patent No. 6458561  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Molteni, Angela  
; APPLICANT: Magnaghi, Paola  
; APPLICANT: Bosotti, Roberta  
; APPLICANT: Scaccheri, Emanuela  
; APPLICANT: Isacchi, Antonella  
; APPLICANT: Hodgson, Dave  
; TITLE OF INVENTION: HUMAN NIM1 KINASE  
; FILE REFERENCE: PC-0009 US  
; CURRENT APPLICATION NUMBER: US/09/523.849  
; CURRENT FILING DATE: 2000-03-13  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PERL Program  
; SEQ ID NO 32  
; LENGTH: 793  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: GenBank Accession No. 6458561 g2052189  
US-09-523-849-32

Query Match 21.0%; Score 691.5; DB 4; Length 793;  
Best Local Similarity 35.6%; Pred. No. 4.3e-51;  
Matches 187; Conservative 76; Mismatches 189; Indels 73; Gaps 14;  
Qy 57 YEFLETIGKTYGKVKARE--SSGRLVAIKSRDKIKDEQDLHRRREIEMSLNHPH 115  
Db 60 YRLKTTKGNFAVKLARHLTLTGREVAIKIDKTL-NPTSLOKLFREVRIMKILNHPN 118  
Qy 116 IIAIHEVFENSCKTIVMEYASRGDLYDIYSERPLSERDARHFFRQIVSALHYCHONGI 175  
Db 119 IVKLFEVETIEKTLVMEYASGVEFDYLAHGRMKKEARAKFRQIVSAVQYCHOKCI 178  
Qy 176 VHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYGVPEV 235  
Db 179 VHRDLKAENLLDADNMKIADFGFSNEFTVGNKLDTPFCGSPPYAAPELFGKYYDGPV 238  
Qy 236 DSWSLGVLLYILVHGTWMPFDGQDHKTLVKQISNGAYREPPEKPSDAC-GLIRWLLMVNPT 294  
Db 239 DVMSGLVILYTLVSGSLPFDGQNLKELRVRVLRKYRIPFYMSTDCENLLKRFVLVLPK 298





QY 402 RPKSSSLKPKGI-----LKKSTSSGE-----VQEDPQELRPV 436  
DB 387 SAPSPSHKQVRSVANSANKORRFSDQAGPAIPTNSYSKKTQSNNAENKRPEEDRESGRKA 446  
QY 437 PDT---PGQVPVAVSLPRKGIKKSQRSGYSSPESESGELLDASDFVSGDPVEQ 493  
DB 447 SSTAKVPASPLP-----GLERKK-----TTPTSTNSVLSTSTN-----RSR 483  
QY 494 KSPQASGLLHRRKILK---NGKFSRTALEGTTPTSTFGSLDQLAGSHPA--RPSRPSGA 549  
DB 484 NSP-----LLERASLGQASIQNGKDS-TAPQ-RVPVAPSPSAHNISSSGGAPDRTNPRGV 536  
QY 550 VSEDSILSS---ESFDQLDLPERL-PETP 574  
DB 537 SSRSTFHAGQLRQVRDQONLPGYVTPASP 565

RESULT 10  
US-10-274-194-2  
; Sequence 2, Application US/10274194  
; Patent No. 6706511  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001306DIV  
; CURRENT APPLICATION NUMBER: US/10/274,194  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 724  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-274-194-2

Query Match 20.6%; Score 680; DB 4; Length 724;  
Best Local Similarity 32.0%; Pred. No. 3.7e-50;  
Matches 201; Conservative 99; Mismatches 195; Indels 134; Gaps 25;

QY 6 LLORPSQAPSASALASESARPLADGLIKSPKPLMKQAVKHHKHHLRHR-----YEFL 60  
DB 11 LNERDTEQPTLGHLD-----KPSKSNMIRGNRSATSDEQPHIGNYRL 56  
QY 61 ETLGKGTGYKKVARE--SSGLVAIKSIRKDKIDQDLHIREIEMSLNHPHIIAI 119  
DB 57 KTIKGNFAKVKLARHILTGKEVAVKIIDKTQL-NSSSLQKLPREVIRIMKVLNHPNIVKL 115  
QY 120 HEVFENSSKIVIVMEYASRGDLYISERPLSRDARHFFRQIVSALHYCHONGIVHRD 179  
DB 116 FEVIETKTLTYLWMEYASGGVEFDYLVNAGRMEKEARAKFRQVSAVQYCHQKFIHRD 175  
QY 180 LKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEDVSW 239  
DB 176 LKAENLLDADNMKIADFGFSNEFTFGNKLDTCGSPPYAAPFLQKGYDGEVDVWS 235  
QY 240 LGVLLYLIVHGTMPDGDQDKHTLVKQISNGAYREPPKPSDAC-GLIRWLLMWNTRATL 298  
DB 236 LGVILYTLVSGSLPFDGQNLKELRERVLRGKYRIPFYMTDCENLLKFLINPSKRGTL 295  
QY 299 EDVASHWVWNGYTTGVGEQALREGHPSGDFGRASMDWLRSSRPLENGAKVCSFF 358  
DB 296 EQIMKDRMNVGH-----EDELKPYVEPLDYKDP-----RTELMSVG-----YT 338  
QY 359 KQHPVGGSTVPGLERQHSLSKSRKENDMA-----QNLOGD-----PAED-TSS 401  
DB 339 REEI-----QDSLQVQRYNEVMATYLLLGYSSELEGDTITLKPRAADLTNS 386  
QY 402 RPKSSSLKPKGI-----LKKSTSSGE-----VQEDPQELRPV 436  
DB 387 SAPSPSHKQVRSVANSANKORRFSDQAGPAIPTNSYSKKTQSNNAENKRPEEDRESGRKA 446

QY 437 PDT---PGQVPVAVSLPRKGIKKSQRSGYSSPESESGELLDASDFVSGDPVEQ 493  
DB 447 SSTAKVPASPLP-----GLERKK-----TTPTSTNSVLSTSTN-----RSR 483  
QY 494 KSPQASGLLHRRKILK---NGKFSRTALEGTTPTSTFGSLDQLAGSHPA--RPSRPSGA 549  
DB 484 NSP-----LLERASLGQASIQNGKDS-TAPQ-RVPVAPSPSAHNISSSGGAPDRTNPRGV 536  
QY 550 VSEDSILSS---ESFDQLDLPERL-PETP 574  
DB 537 SSRSTFHAGQLRQVRDQONLPGYVTPASP 565

RESULT 11  
US-09-984-890-4  
; Sequence 4, Application US/09984890  
; Patent No. 6492156  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001306  
; CURRENT APPLICATION NUMBER: US/09/984,890  
; CURRENT FILING DATE: 2001-10-31  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 722  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-984-890-4

Query Match 20.8%; Score 677.5; DB 4; Length 722;  
Best Local Similarity 32.9%; Pred. No. 6.1e-50;  
Matches 201; Conservative 92; Mismatches 203; Indels 115; Gaps 24;

QY 6 LLORPSQAPSASALASESARPLADGLIKSPKPLMKQAVKHHKHHLRHR-----YEFL 60  
DB 11 LNERDTEQPTLGHLD-----KPSKSNMIRGNRSATSDEQPHIGNYRL 56  
QY 61 ETLGKGTGYKKVARE--SSGLVAIKSIRKDKIDQDLHIREIEMSLNHPHIIAI 119  
DB 57 KTIKGNFAKVKLARHILTGKEVAVKIIDKTQL-NSSSLQKLPREVIRIMKVLNHPNIVKL 115  
QY 120 HEVFENSSKIVIVMEYASRGDLYISERPLSRDARHFFRQIVSALHYCHONGIVHRD 179  
DB 116 FEVIETKTLTYLWMEYASGGVEFDYLVNAGRMEKEARAKFRQIVSAVQYCHQKFIHRD 175  
QY 180 LKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPEDVSW 239  
DB 176 LKAENLLDADNMKIADFGFSNEFTFGNKLDTCGSPPYAAPFLQKGYDGEVDVWS 235  
QY 240 LGVLLYLIVHGTMPDGDQDKHTLVKQISNGAYREPPKPSDAC-GLIRWLLMWNTRATL 298  
DB 236 LGVILYTLVSGSLPFDGQNLKELRERVLRGKYRIPFYMTDCENLLKFLINPSKRGTL 295  
QY 299 EDVASHWVWNGYTTGVGEQALREGHPSGDFGRASMDWLRSSRPLENGAKVCSFF 358  
DB 296 EQIMKDRMNVGH-----EDELKPYVEPLDYKDP-----RTELMSVG-----YT 338  
QY 359 KQHPVGGSTVPGLERQHSLSKSRKENDMA-----QNLOGD-----PAED-TS 400  
DB 339 REEI-----QDSLQVQRYNEVMATYLLLGYSSELEGDTITLKPRAADLTNS 386  
QY 401 SRPKSKSLKPKGIKKSSTSGVEQEDPQELRPVDPDTPQVPAVSLPRKGIKKS 460  
DB 387 SAPSPSH-----KVORSVSA-----NPQRR-----SSDQAVPAI---PTSNSYSKKT 426  
QY 461 QRSQYSSPESESGELLDASDFVSGDP-VEQK-----SPOASGL-----LL 503  
DB 427 QSNNAENKRPEETGRKASSTAKVPASPLGLRDKTTPTSTNSVLSTSTNSRNSPL 486

Qy 504 HRKGLK---NGKFSRTALEGTTPTSGLDQLASSHPAA-RPSRSGAVSEDSILSS- 558  
Db 487 DRASLGQASTQNGKDS-TAPO-RVFPVAPSAHNISSSSGAPDRTNFRPGVSSRSTFFHAGQ 544  
Qy 559 --ESFDQDLDP 567  
Db 545 LRQVRDQONLP 555

RESULT 12  
US-10-274-194-4  
; Sequence 4, Application US/10274194  
; Patent No. 6706511  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001306DIV  
; CURRENT APPLICATION NUMBER: US/10/274,194  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 722  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-274-194-4

Query Match 20.6%; Score 677.5; DB 4; Length 722;  
Best Local Similarity 32.9%; Pred. No. 6.1e-50;  
Matches 201; Conservative 92; Mismatches 203; Indels 115; Gaps 24;

Qy 6 LLQRPQAPSASALASESARPLADGLIKSPKPLMKQAVKHHHKLNRH-----YEFL 60  
Db 11 LNERDTEQPTLGHLS-----KPSKSNWLRGRNSATGADQPHIGNVRL 56

Qy 61 ETLGKTYGVKKARE-SSGRLVAIKSIRKDKDEQDLHRRREIEMSLNHPHIIAI 119  
Db 57 KTIGKGNFAVKLARHILTGKAVAKIIDKTQL-NSSSLQKLPREVIMKVLNHPNIVKL 115

Qy 120 HEVPENSKIVIVMEYASRGDLYDYSERPLSRDARHFRQIVSALHYCHQNGI 179  
Db 116 FEVIEETKTYLVMEYASGGEVFDYLVAGHMKKEARAKFRQIVSAVQYCHQKFI 175

Qy 180 LKLENIILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPVDWS 239  
Db 176 LKAENLLDADNMNIKIADFGSNEFTFGNKLDTFCGSPPYAAPLFGKKYDGEVDWS 235

Qy 240 LGVLLYLIVHGTWPFQGDHKTLLVKQISNGAYREPPKPSDAC-GLIRWLLMVPTRATL 298  
Db 236 LGVILYTLVSGSLPFDGQNLKELRERVLRGYRIPFYMWSTDCENLLKFLILNPSKGT 295

Qy 299 EDVASHWVWNGYTTGVGEQALREGGHPGDFGRASMDWLRSSRPLENGAKVCSFF 358  
Db 296 EQIMKDRMNVGH-----EDELKPYVPELPDYKDP-----RRTLESMVG-----YT 338

Qy 359 KOHVPGGGSTVGLERQHSILKSKSKENDMA-----QNLOGD-----PAED--TS 400  
Db 339 REEL-----QDSLVGRQYNEVMATYLLGYKSELEGDTTLKPRPSADLTNS 386

Qy 401 SRPGKSLKLPKGLKKKSSSTSGEVEDPQELRPVDPDTPGQVPVAVSLLPRKGLKKSR 460  
Db 387 SAPSPSH-----KVQSVSA-----NPKQR-----SSDQAVAI--PTSNSYSKKT 426

Qy 461 QRESGYSSPEPSSEGLLDASDVFSVGGDP-VEQK-----SPQASGL-----LL 503  
Db 427 QSNNAENKRPETEGRKASSTAKVPASPLPGLDRKKTTPSTNSVLSTSTNSRNSPLL 486

Qy 504 HRKGLK---NGKFSRTALEGTTPTSGLDQLASSHPAA-RPSRSGAVSEDSILSS- 558  
Db 487 DRASLGQASTQNGKDS-TAPO-RVFPVAPSAHNISSSSGAPDRTNFRPGVSSRSTFFHAGQ 544

Qy 559 --ESFDQDLDP 567  
Db 545 LRQVRDQONLP 555

RESULT 13  
US-08-817-832B-31  
; Sequence 31, Application US/08817832B  
; Patent No. 6579691  
; GENERAL INFORMATION:  
; APPLICANT: MANDELKOW, Eckhard, et al.  
; TITLE OF INVENTION: No. 6579691el Protein Kinase (NPK-110)  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 S. Wacker Drive, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: US  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/817,832B  
; FILING DATE: 28-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/EP95/04258  
; FILING DATE: 30-OCT-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 94 11 7122.5  
; FILING DATE: 28-OCT-1994  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 779 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-817-832B-31

Query Match 20.5%; Score 675; DB 4; Length 779;  
Best Local Similarity 35.4%; Pred. No. 1.1e-49;  
Matches 186; Conservative 75; Mismatches 191; Indels 74; Gaps 15;

Qy 57 YEFLETGKTYGVKKARE-SSGRLVAIKSIRKDKDEQDLHRRREIEMSLNHPH 115  
Db 46 YRLQKTIGKGNFAVKLARHVLTGREAVAKIIDKTQL-NPTSLQKLPREVIMKILNHPN 104

Qy 116 IIAIHEVPENSKIVIVMEYASRGDLYDYSERPLSRDARHFRQIVSALHYCHQNGI 175  
Db 105 IVKLFVEIETKTYLVMEYASGGEVFDYLVAGHMKKEARAKFRQIVSAVQYCHQKCI 164

Qy 176 VHRDLKLENIILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPVEV 235  
Db 165 VHRDLKLENIILLDANGNIKIADFGSNEFTVGNNKLDTCGSPPYAAPLFGKKYDGEV 224

Qy 236 DWSLGLVLLYLIVHGTWPFQGDHKTLLVKQISNGAYREPPKPSDAC-GLIRWLLMVPNT 293  
Db 225 DVMSLGVILYTLVSGSLPFDGQNLKELRERVLRGYRIPFYMWSTDCENLLKFLILNPI 284

Qy 294 RRATLDRVASHWVWNGYTTGVGEQALREGGHPGDFGRASMDWLRSSRPILLE-NGA 352  
Db 285 KRGSBOIMKDRMNVGH-----EEBELKPYSEPELDLNDAKRIDINMTWTFGARDEINDA 339

Qy 353 KVCSEFFKQ-----HVPGGGSTVGLERQHSILKSKSKENDMAQNLOGDPAEDT 399  
Db 340 LVQSKYDEVWATYLLGRKPPEEGGESLSSG-----NLCSRPSSDLNNTLOSFAH-- 393

Qy 400 SSRPGKSSLLKPKGLKKKSSSTSGEVEDPQELRPVDPDTPGQVPVAVSLLPR-KGILK 457

Db 394 -----LKQRTISANQKQ-----RRFSDHAGSIPPAAVSYTKRPQANSV 432  
Qy 458 KSRQRE-----SGYSSP--PSESSEGLLDASDVFSVGDVPEQKSPQ 497  
Db 433 ESEQKEWMDKTARLGGTTVGSSEVTSPLVGPDRKSSAGPSNNVYSGGSMTRNTY 492  
Qy 498 ASGLLLHRKGLKNGKFSRTALEGTTPTSTFGSLDQLASSHPPAARP 543  
Db 493 VCERSTURYAALONGRSSLSLTEMSSSTGS--TVASAGPSARP 536  
RESULT 14  
US-08-817-832B-32  
; Sequence 32, Application US/08817832B  
; Patent No. 6579691  
; GENERAL INFORMATION:  
; APPLICANT: MANDELKOW, Eckhard, et al.  
; TITLE OF INVENTION: NO. 6579691e1 Protein Kinase (NPK-110)  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 S. Wacker Drive, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: US  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/817,832B  
; FILING DATE: 28-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/EP95/04258  
; FILING DATE: 30-OCT-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 94 11 7122.5  
; FILING DATE: 28-OCT-1994  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 722 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-817-832B-32

Query Match 20.5%; Score 673.5; DB 4; Length 722;  
Best Local Similarity 31.3%; Pred. No. 1.4e-49;  
Matches 208; Conservative 88; Mismatches 212; Indels 157; Gaps 25;  
Qy 6 LLQRPQASASALASAPRADGLIKSPKLMKQAVKRHHKHLRHR-----YEFL 60  
Db 11 LNERDTEQPTGLHDS-----KPSKSNLMRGNSATSADQPHIGNYRL 56  
Qy 61 ETLGKGYGVKKARE--SSGRLVAIKIRKDKIKDEQDLHIREIIMSLNPHIIAI 119  
Db 57 KTIKGKGNFAKVKLARHLTGKEVAVKIDKTLQ--NSSSLQKLFREVRIMKVLNHPNIVKL 115  
Qy 120 HEVPENSSKIVIMVEYASRGDLYDYSERPLSRDARHFRQIVSALHYCHQNGIVHRD 179  
Db 116 FEVIETKTLVLMVEYASGSEVFDYLVHAGMKKEARAFQIVSAVQVCHHKFIVHRD 175  
Qy 180 LKLENILLDANGNIKIADFGLSNLYHKGFELQTCGSPLYASPIVNGKPYVGEVDWS 239  
Db 176 LKAENLLLDAMNIIADFGFSNEFTFGNKLDTFCGSPPYAAPELFGKKYDGEVDVWS 235  
Qy 240 LGVLLYILVHGTMPFDQDQDKTLVKQISNGAYRPPKPSDAC--GLIRWLLMNPTRATL 298  
Db 236 LGVILYTLVSGSLPFDQGNLKERLVRNGKIRIFPFYMTDCENLLKFKLILNPSKRGTL 295

Qy 299 EDVAHSHWVN-----KGYYTTG----- 314  
Db 296 EQIMKORWMMVGHEDDELKPYVEPLPDYKDPRTTELMMVSMGYTREETQDSLGVQRYNEVM 355  
Qy 315 -----VGEQBALREGG-----HPSGDFGRASMDWLRSSRPPLLENGAKVCGPFFKQHPV 363  
Db 356 ATVLLGLGYSSELEGGDTITLKPRPSADLTNSSAPSPSHKQVRSVSNP-----KQR-R 407  
Qy 364 GGGSVTPGLERQHSLSKKSRKENDMAQNLQGDPAEDTSSRPGKSLKLP-----KGILKKKS 419  
Db 408 SSDQAVPAIPTNSYSKKTQSN-----NAENKRPREETGRKASSTAKVPASPLPGLDRKKT 463  
Qy 420 -----STSGEVEQEDPOELRPVDPDPGPVPAVSLPRKILKKSRQRESGYSS 469  
Db 464 TPTPSTNSVLSTSTNRSNP-----LLDRASLGQASIQ--NGKDS 503  
Qy 470 -----PEPSESGLLDASDVFSVGDVPEQKSPQ--ASGLLHHRKGLKNGKFSRTALEG 522  
Db 504 APQKVPVAPSAHNISS-----SGAPDRTNFRGVSSRSSTFHA-----GQLRQVRDQ 552  
Qy 523 TPTSTFGSLDQLASSHPPAARPSPRGAVSEDSILS--SESFQDLDLPERLPETPLRCVSV 580  
Db 553 NLP--FGVTPASPSPSHGSHQGRG--PSG-----SIFSKFTSKFVRRLNE--PESKDR 598  
Qy 581 VDNLR 585  
Db 599 VETLR 603  
RESULT 15  
US-09-523-849-35  
; Sequence 35, Application US/09523849  
; Patent No. 6458561  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Molteni, Angela  
; APPLICANT: Magnaghi, Paola  
; APPLICANT: Bosetti, Roberta  
; APPLICANT: Scacchi, Emanuela  
; APPLICANT: Isacchi, Antonella  
; APPLICANT: Hodgson, Dave  
; TITLE OF INVENTION: HUMAN NIM1 KINASE  
; FILE REFERENCE: PC-0009 US  
; CURRENT APPLICATION NUMBER: US/09/523,849  
; CURRENT FILING DATE: 2000-03-13  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PERL Program  
; SEQ ID NO 35  
; LENGTH: 604  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: GenBank Accession No. 6458561 g2564680  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 52, 291, 354  
; OTHER INFORMATION: unknown or other  
; US-09-523-849-35  
Query Match 19.8%; Score 653; DB 4; Length 604;  
Best Local Similarity 33.0%; Pred. No. 6.2e-48;  
Matches 190; Conservative 76; Mismatches 206; Indels 104; Gaps 19;  
Qy 57 YEFLETGKGYGVKKARE--SSGRLVAIKIRKDKIKD--EQDLHIREIIMSLN 113  
Db 63 YKIITKGNFAKVKLARHLTGREVAIKIDKLTALNTIARQKLY---REVNMKKLNH 119  
Qy 114 PHIIATHEVPENSSKIVIMVEYASRGDLYDYSERPLSRDARHFRQIVSALHYCHQNG 173  
Db 120 FNIVRLQVIESRTLYLVMEYVSGSEVFDYLVHAGMKKEARAFQIVSAVQVCHHKF 179  
Qy 174 GIVHRDLKLENILLDANGNIKIADFGLSNLYHKGFELQTCGSPLYASPIVNGKPYVGP 233

```
Db 180 SIHRDLKAENLLDQOMKLIADFGFTTPEPKAPLETFCGSPPYAAPLFGKGYSGP 239
Qy 234 EVDSWSLGVLLYILVHGTMPDPDGDHKTLYKQISNGAYREPPKPSDAC-GLIRWLLAVNP 292
Db 240 EVDSWSLGVLLYILVHGTMPDPDGDHKTLYKQISNGAYREPPKPSDAC-GLIRWLLAVNP 299
Qy 293 TRRATLEDAVASHWVNNNGYTTGVGEQALREGGHPSGDFGRASMAWLRSSRPILLENGA 352
Db 300 TORTSLSAVWADRWNNNGYEGNG-LRPFQEKPMDLHDVNRLLSLLSNMGHKPRDVXQSLK 358
Qy 353 K-----VCSFFKOHVPGGSGTVFGLERQHSLSKSRKENDMAQNLQGDPAEDTSSRPCKS 406
Db 359 NQKFDIYCAVMLLDVAKPRSTACSEKSGSSFR-----ETPTAMPGSS 401
Qy 407 SLKLPKG-----ILKKSSTSSGEVQEDP-----QELRPVDPDTPGQVPAVSL 449
Db 402 RIPVIAAPNVTISQVTFALDKSTPNRPGATSIKMAPRIANALTPLPLTPP----- 453
Qy 450 LPRKGIL----KKSRORESGYSSPEPS--ESGELLDDASDVFSVSGDPVEQKSPQASGLLL 503
Db 454 -PKYICCSASKAANPRSRPSSITPQSAMPKG-----VGSFVDVKT-----TLL 497
Qy 504 HRKGILKNGKFSRTALEGTTPTFTGSLDQLASHPAARP-SRPSGAVSEDSILSSEFSD 562
Db 498 SAQRKLAHVNHKLT-----SASHQIRSPITQSSSQASECTRTPPTPHTFE 539
Qy 563 QLDLPERLPETPLRGCVSDNLRGLEOPPS-EGLKR 597
Db 540 MLD-----STSTPLKVLKLVAS--NSQTPPSTENINR 569
```

Search completed: January 22, 2005, 06:44:54  
Job time : 43.6867 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 06:29:16 ; Search time 38.5452 Seconds  
(without alignments)  
1575.104 Million cell updates/sec

Title: US-09-980-464-11  
Perfect score: 3293  
Sequence: 1 MESVALLQRPQAPSASALA.....DCOEVTAAVRQALGICSKLS 631

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	888	27.0	1558	2 T29253	hypothetical prote
2	694.5	21.1	713	2 S27966	probable serine/th
3	689	20.9	745	2 G01025	serine/threonine p
4	686.5	20.8	774	2 I48609	probable serine/th
5	669	20.3	1192	2 T18611	probable serine/th
6	667	20.3	1246	2 G89287	protein H39823.1
7	664	20.2	1398	2 T13741	hypothetical prote
8	649.5	19.7	512	2 T52633	serine/threonine-s
9	648.5	19.7	633	1 A26030	serine/threonine-s
10	648	19.7	798	2 JC7500	gik protein - chic
11	641	19.5	504	2 T10449	probable serine/th
12	639.5	19.4	512	1 JC1446	serine/threonine-s
13	631	19.2	511	1 A56009	serine/threonine-s
14	630.5	19.1	552	1 S51025	hydroxymethylglut
15	627.5	19.1	602	2 S72513	FG2 protein - yea
16	623.5	18.9	552	1 A53928	hydroxymethylglut
17	622	18.9	562	2 T29858	hypothetical prote
18	621	18.9	472	2 B90100	SNF-related kinase
19	620	18.8	504	2 T07415	probable serine/th
20	612	18.6	481	2 I49072	protein kinase - m
21	610	18.5	473	1 S59941	serine/threonine-s
22	599	18.2	512	2 T07788	probable serine/th
23	596.5	18.1	887	2 T20941	hypothetical prote
24	594.5	18.1	576	2 T41587	probable carbon ca
25	587.5	17.8	622	1 S44859	serine/threonine-s
26	583.5	17.7	891	2 A38903	protein kinase 1 -
27	581.5	17.7	891	2 T40503	protein kinase kin
28	579	17.6	502	2 T04145	serine/threonine p
29	570	17.3	502	1 A41361	serine/threonine-s

SNF1-related prote  
serine/threonine-s  
serine/threonine-s  
serine/threonine-s  
probable protein k  
serine/threonine-s  
p69Bg3 protein - A  
GIN4 protein - yea  
protein kinase KIN  
probable serine/th  
probable serine/th  
probable serine/th  
hypothetical prote  
SNF1-related prote  
probable protein k

ALIGNMENTS

RESULT 1

T29253  
hypothetical protein B0496.3 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T29253  
R;Murray, J.; Le, T. T.  
Submitted to the EMBL Data Library, May 1996  
A;Description: The sequence of C. elegans cosmid B0496.  
A;Reference number: Z20596  
A;Accession: T29253  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1558 <MUR>  
A;Cross-references: EMBL:U58749; PIDN:AAB00636.1; GSPDB:GN00022; CESP:B0496.3  
A;Experimental source: strain Bristol N2; clone B0496  
C;Genetics:  
A;Gene: CESP:B0496.3  
A;Map position: 4  
A;Introns: 55/1; 100/3; 141/1; 226/3; 298/1; 468/3; 502/2; 609/3; 656/3; 795/3; 843/2; 9

Query Match	27.0%;	Score	888;	DB 2;	Length	1558;
Best Local Similarity	37.0%;	Pred. No.	2.1e-28;			
Matches	230;	Conservative	73;	Mismatches	214;	Indels 104; Gaps 15;
QY	54	RHRYELETGLGTYGKVKKARESS-GRUVAKISRKDKIKDBODLLHRRREIEMSSLN	112			
DB	72	KHREITKLGSGTYGKVSLAYDHHKFDREAVAVKLKKSADLVIRIRREIRIMSALN	131			
QY	113	HPHIIAHEVFENSKITIVMEYASRGDLVDYISERPLSERDARHFFROIYSALHYCHQ	172			
DB	132	HPNIIQIYEFENKDKIILVMEYSSGGELYDYVSRGCSLPEAEARRIFQITSAVLCHK	191			
QY	173	NGIVHRDLKENILLDANGNIKIADPGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYVG	232			
DB	192	HRVAHRDLKENILLDQNNAKIADPGLSNYPADKLLTTFCCGSPLYASPEIINGTYPGK	251			
QY	233	PEVDSWSGLVLLYILVHGTMPPFDGQDHKTLLVKQISNGAYREPPKPSDACGLIRWLLMVP	292			
DB	252	PEVDCWSLGLITLLVYGSGMPFDGRDNRMVRQIKRGAYPEPETPTASTASMLIRMLRVNP	311			
QY	293	TREATLEDVASHWVWNGYTTGV-----GQREALREG	324			
DB	312	ERRATIFDIASHWMLNLENMNPVQELPENQIIDHTPLTERBETMIVQDLADSDQVMEF	371			
QY	325	GHPSGDGFRASMDWLR-----SSRPLENGAKVCSFPKQHPVGGSTV	369			
DB	372	GHLSSTRKIEDFRIRREKAESEFNDSVPKPPKARKTDELTKGISKEQPEEMKSAEKL	431			
QY	370	PGLERQHS-----LKKSRK-ENDMAQNLQGDPAEDTSRPGKSSLLKPKGLIKKK	418			
DB	432	RGVKEKEKPKVVDNDPLERLQIENRLQNKQKDEAAKTSARVETVKLK-----EVK	485			

```
Qy 419 STSSGVEQSDPQELRPVDTPGQVPVAVSLLPRKGLKKSQRQESGYSPSPESGEL 478
Db 486 KEKSPQEPDPPKAR-----GTSKPADSRAP-SFVVDKRPDT-----SEPERPRTRPH 534
Qy 479 LDASDFVFSQDPVEQKSPQASGLLHLRKGILKLN--GKFSRTALEGTTPTSTFGSLDQLAS 536
Db 535 LTASAVRIETDSLNLMLNQV--LEQMEKGPVNLNIARIKAHPLYDTRPVMVKELLESIIA 592
Qy 537 SHPAARSPRSGAVSDSILSSSFDQL-----DLPERLPETPLRGCVSDNLRGLEPPS 592
Db 593 AQPEPVQKQTSKVVEQQTFSRQNTLTRKKKEDPLEEPIEVP-----SPS 637
Qy 593 EGLK-RWV-----QESLG 604
Db 638 RKMKEPWHVSVECMKNESVG 658

RESULT 2
S27966
probable serine/threonine-specific protein kinase (EC 2.7.1.1) - human
N;Alternate names: protein p78
C;Species: Homo sapiens (man)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 16-Aug-2004
C;Accession: S27966
R;Maheshwari, K.K.; Som, S.; Parsa, I.
submitted to the EMBL Data Library, January 1992
A;Description: Sequence of a cDNA encoding 78kD marker protein lost in chemically induced
A;Reference number: S27966
A;Accession: S27966
A;Molecule type: mRNA
A;Residues: 1-713 <MAH>
A;Cross-references: UNIPROT:P27448; EMBL:M80359; NID:g189511; PID:AAAS9991.1; PID:g1895
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F;54-307/Domain: protein kinase homology <KIN>
F;62-70/Region: protein kinase ATP-binding motif

Query Match 21.18; Score 694.5; DB 2; Length 713;
Best Local Similarity 34.5%; Pred. No. 4.9e-21;
Matches 191; Conservative 75; Mismatches 171; Indels 117; Gaps 17;

Qy 57 YEFLTLGKGTGKVKKARE--SSRLVAIKSIRKDKIKDEQLLHRIREIEIMSSLNHPH 115
Db 56 YRLKTIKGNFAKVKLARHLITGREVAIKIIDTKQL-NPTSLOKLFREVRIMKVLNHPN 114
Qy 116 IIAIHEVFENSXKIVIVMEYASRGDLYDISERPLSRDARHFRRQIVSALHYCHQNGI 175
Db 115 IVKLFEVETQKTLYLIMEYASGGKVPDYLVAHGRMEKEARSFRQIVSAVQYCHQKRI 174
Qy 176 VHRDLKLENILLDANGNIKIADPGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYVGPV 235
Db 175 VHRDLKAENLLDADNMKIADFGFSNEFTVGGKLDITFCGSPPYAAPLFPQKKYDGPV 234
Qy 236 DWSLGLVLLYLTHGTMPPFDGQDHKTLVKQISNGAYREPPKPSDAC--GLIRWLLMVNPT 294
Db 235 DVNSLGVLLYTLVSGSLPFDGQNLKELRERVLKGIPIPFYMTDCENLKKFLVLPNIK 294
Qy 295 RATLEDVASHWVNWGYTTGVGEQALREGGHPGDFGRASMDLRRSRRLLENGAKV 354
Db 295 RGTLEQIMKDRWINAGH-----EBDELKPFVEPELDISDQKRIDIM-----V 336
Qy 355 CSFQKHVPGGSTVPGLEERQHSLLKSKRKENDMAQNL-----QCPAEDTSRRPKSS 407
Db 337 GMGYSQSE-----ETQESLKMKYDEITATYLLGRKSESEVRPSSDLNNSTQGS- 384
Qy 408 LKLPKGLIKKKSSTSSGVEQDPELRVPDTPGQVPVAVSLLPR----- 452
Db 385 ---PHHKVQSVSS-----QKRRYSDHAGPGIPSVVAYPKRSQTSADSLKEDG 433
Qy 453 -----KGILKKSQRQESGYSPSPESGELLDASDFVFSQDPVEQKSPQAS 499
Db 434 ISSRKSTGSAVGKGIAAPSM--LGNASPNKADIPERKKSSTV-----PSSNTAS 483
```

```
Qy 500 GLLLRHK-----GILKLNKKSRTALEGTTPTSTFGSLDQLASSH---PAARP 543
Db 484 GGMTRRTVYVCSBRTTDRHSVIQ-NGKENSTIPDQRT-----VASTHSSAATP 534
Qy 544 SR---PSGAVSEDS 554
Db 535 DRIRFPRGTASRST 548

RESULT 3
G01025
serine/threonine protein kinase - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 16-Aug-2004
C;Accession: G01025
R;Navarro, E.
submitted to the EMBL Data Library, April 1996
A;Reference number: H00564
A;Accession: G01025
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-745 <NAV>
A;Cross-references: UNIPROT:Q15524; UNIPROT:Q96HB3; EMBL:X97630; NID:g1310674
C;Superfamily: protein kinase homology
F;18-271/Domain: protein kinase homology <KIN>

Query Match 20.9%; Score 689; DB 2; Length 745;
Best Local Similarity 33.2%; Pred. No. 8.4e-21;
Matches 189; Conservative 92; Mismatches 176; Indels 112; Gaps 20;

Qy 57 YEFLTLGKGTGKVKKARE--SSRLVAIKSIRKDKIKDEQLLHRIREIEIMSSLNHPH 115
Db 20 YRLKTIKGNFAKVKLARHLITGREVAIKIIDTKQL-NSSSLOKLFREVRIMKVLNHPN 78
Qy 116 IIAIHEVFENSXKIVIVMEYASRGDLYDISERPLSRDARHFRRQIVSALHYCHQNGI 175
Db 79 IVKLFEVETQKTLYLIMEYASGGVFDYLVAHGRMEKEARAKFRQIVSAVQYCHQKFI 138
Qy 176 VHRDLKLENILLDANGNIKIADPGLSNLYHKGFLOTFCGSPLYASPEIVNGKPYVGPV 235
Db 139 VHRDLKAENLLDADNMKIADFGFSNEFTVGGKLDITFCGSPPYAAPLFPQKKYDGPV 198
Qy 236 DWSLGLVLLYLTHGTMPPFDGQDHKTLVKQISNGAYREPPKPSDAC--GLIRWLLMVNPT 294
Db 199 DVNSLGVLLYTLVSGSLPFDGQNLKELRERVLKGIPIPFYMTDCENLKKFLVLPNIK 258
Qy 295 RATLEDVASHWVNWGYTTGVGEQALREGGHPGDFGRASMDLRRSRRLLENGAKV 354
Db 259 RGTLEQIMKDRWINVGH-----EDDELKPYVEPLPDYKDP-----RRTLMVSMG--- 303
Qy 355 CSFQKHVPGGSTVPGLEERQHSLLKSKRKENDMA-----ONLOGD-----PAED 398
Db 304 --YTREEI-----QDSLVGQRYNEVMATYLLGLYKSELEGGDTITLKPRPSAD 349
Qy 399 -TSRRPKKSLKPKGI-----LKKKSTSSSG----VQEDPQE 432
Db 350 LTNSSAQFPHKQVRSVANSAPKORRSDQAGPAIPTSNYSKKTQSNNAENKPEEDRES 409
Qy 433 LRVPDPT---PGQVPVAVSLLPRKGLKKSQRQESGYSPSPESGELLDASDFVFSGD 489
Db 410 GRKASSTAKVPASPLP-----GLERKK-----TTFTPTNSVLSTSTN----- 447
Qy 490 PVQKSPQASGLLHLRKGILKLN--GKFSRTALEGTTPTSTFGSLDQLASSHPAARSRP 546
Db 448 -RSRNSP-----LLERASLGQASIQNGKDSLT-MPGSRASTASAGVASAARPRQHOKSM 500
Qy 547 SGAV---SEDSILSSSFDQLDPLRPE 572
Db 501 SGSVHPNKASGLPPTESNCEVPRPSTAPQ 529

RESULT 4
```



J148609

probable serine/threonine-specific protein kinase (EC 2.7.1.1-) kem - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 16-Aug-2004  
C:Accession: J148609; S31333  
R:Ingalls, J.D.; Lee, M.; Hill, R.E.  
Mamm. Genome 4, 401-403, 1993  
A:Title: Emk, a protein kinase with homologs in yeast maps to mouse chromosome 19.  
A:Reference number: J148609; MUID:93364122; PMID:8358177  
A:Accession: J148609  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-774 <ING1>  
A:Cross-references: UNIPROT:Q05512; EMBL:X70764; NID:G57919; PIDN:CAA50040.1; PID:G57920  
R:Ingalls, J.D.; Lee, M.; Hill, R.E.  
submitted to the EMBL Data Library, January 1993  
A:Description: A novel protein kinase with homologues in yeast maps to mouse chromosome  
A:Reference number: S31333  
A:Accession: S31333  
A:Molecule type: mRNA  
A:Residues: 1-698, 'K', 700-702, 'GRGLPTAAKASCTSGTCYCAGMAHQART', 731-774 <ING2>  
A:Cross-references: EMBL:X70764  
C:Genetics:  
A:Gene: emk  
C:Superfamily: protein kinase homology  
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
F:51-304/Domain: protein kinase homology <KIN>  
F:59-67/Region: protein kinase ATP-binding motif

Query Match 20.8%; Score 686.5; DB 2; Length 774;  
Best Local Similarity 32.6%; Pred. No. 1.1e-20;  
Matches 199; Conservative 90; Mismatches 218; Indels 103; Gaps 20;  
6 LLQRPQAPASALASASARPLADGLIKSPKPLMKQAVKHHKHLRHR-----YEFL 60  
11 LNERDTQPTGLHLD-----KPSKSNLGRNSATSADQPHIGNVRL 56  
61 ETLLGKGTGVKKARE-SSGRLVAIKSIRKDKIDEQDLHRRIRIEMSLNPHIIAI 119  
57 KTIGKGNFAKVLARHILTGKVAVKIIDKTQL-NSSSLQKLFREVRIMKVLNHPNIVKL 115  
120 HEVFENSKIVIVMEYASRGDLYIISERPLSRDARHFRQIVSAHYCHQNGIVHRD 179  
116 FEVIETETKLYLVMYASGGVEFDYLVAGHGMKEKARAKFRQIVLVHVOYCHQKFIHRD 175  
180 LKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPVGVGVDSWS 239  
176 LKAENLLDAMNIIKIADFGSNFTFGNKLDTFCGSPPYAAPLFGQKIDGPEVDVWS 235  
240 LGVLLYILVHGTMPFDGQDHKTLVKQISNGAYREPFPKPSDAC-GLIRWLLMVNPTRRATL 298  
236 LGVILYTLVSGSLPFDQNLKEURVLRGKYRIPFYMWSTDCENLLKKFLILNPSKGTLL 295  
299 EDVASHVWVNGYTTGVGEQALREGGHPGDFGRASMDWLRRSSRPILLENGAKVCSFF 358  
296 EQIMKDRWN-----VGHED-----DELKPYVEPLLTGPD----- 325  
359 KQHVPGGSGTVPGI-----ERQHSCLKSRKENDMAQNL-----QCPAEDTSSRCKSLKL 410  
326 RDRVDG-----VNGLHTEIQQSLVGQRYNEVMATYLLGLYKSEPEGDTITLKPSPADL 381  
411 PKGILKKKSTSGEVEDQELRPVDPDTPQGPVAVSLLPRKILKKSRQESGYSSP 470  
382 TNSSAPSPSHKQVRSVANPKQR-----SSDQAVPAI-----PTNSYSKKTQSNNAENKRP 434  
471 EPESGELLDAODVYVSGDP-VQOK-----SPQASGL-----LHRRGILKL-- 511  
435 EEETGRKASSTAKVPASPLPGLDRKKTTPAPSTNSVLSTSTNSRNSPPLDRASLGQASI 494  
512 -NGKFSRTALEGTTPTFGSILDQASSHP-----AARPPSPCAVSEDSILSESF 561  
495 QNGKDSLT-MPGRASSTASAAVSAARPROHQKMSASVHPNPKASGLPTES--NCEVP 551

Qy 562 DQLDLPERLP 571  
Db 552 RPSTAPQRP 561

## RESULT 5

T18611

probable serine/threonine-specific protein kinase (EC 2.7.1.1-), long splice form - Caenorhadinus  
N:Contains: probable serine/threonine kinase, short splice form  
C:Species: Caenorhadinus elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T18611; T18610; T23144; T23143  
R:McMurray, A.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z18997  
A:Accession: T18611  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1192 <WIL1>  
A:Cross-references: UNIPROT:Q9TW45; EMBL:Z81027; PIDN:CAB54179.1; GSPDB:GN00023; CESP:H39E23.1b  
A:Experimental source: clone AH10  
A:Accession: T18610  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-487, 536-1192 <WIL2>  
A:Cross-references: EMBL:Z81027; PIDN:CAB54178.1; GSPDB:GN00023; CESP:H39E23.1b  
A:Experimental source: clone AH10

R:McMurray, A.

submitted to the EMBL Data Library, June 1997

A:Reference number: Z19696  
A:Accession: T23144  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1192 <WIL3>  
A:Cross-references: EMBL:Z96102; PIDN:CAB54263.1; GSPDB:GN00023; CESP:H39E23.1a  
A:Experimental source: clone H39E23  
A:Accession: T23143  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-487, 536-1192 <WIL4>  
A:Cross-references: EMBL:Z96102; PIDN:CAB54262.1; GSPDB:GN00023; CESP:H39E23.1b  
A:Experimental source: clone H39E23

C:Genetics:

A:Gene: CESP:H39E23.1a; CESP:H39E23.1b

A:Map position: 5

A:Insertions: 10/3; 33/3; 67/3; 139/2; 183/3; 451/3; 487/3; 535/3; 631/1; 825/2; 914/3; 992/3  
C:Keywords: alternative splicing; ATP; phosphotransferase; serine/threonine-specific protein kinase, long splice form #8;  
F:1-1192/Product: probable serine/threonine-specific protein kinase, long splice form #8;  
F:1-487, 536-1192/Product: probable serine/threonine-specific protein kinase, short splice form #8

Query Match 20.3%; Score 669; DB 2; Length 1192;

Best Local Similarity 32.4%; Pred. No. 8.2e-20;

Matches 187; Conservative 97; Mismatches 218; Indels 76; Gaps 16;

Qy 11 SOAPASALASASARPLADG-----LIKSPKPLMKQAVKHHKHLRHRYE 58  
Db 113 SGAPAASS-GGSSARYSSSGRSHPTSGSSSSSHARSTGSCMSRSSRRAARRNDQDVHVCKYK 171  
Qy 59 FLTLGKGTGVKKARE-SSGRLVAIKSIRKDKIDEQDLHRRIRIEMSLNPHII 117  
Db 172 LLKTIGKGNFAKVLARHILTGKVAIKIIDKTAL-NPSSLQKLFREVRIMKVLNHPNIV 230  
Qy 118 AIHEVFENSKIVIVMEYASRGDLYIISERPLSRDARHFRQIVSAHYCHQNGIVH 177  
Db 231 KLYQVMEETQTLVILVYASGGVEFDYLVAGHGMKEKARAKFRQIVSAVOYLHRSKNIIH 290  
Qy 178 RDLKLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPVGVGVDS 237  
Db 291 RDLKAENLLDAMNIIKIADFGSNFTFGNKLDTFCGSPPYAAPLFGQKIDGPEVDV 350  
Qy 238 WSLGVLLYILVHGTMPFDGQDHKTLVKQISNGAYREPFPKPSDAC-GLIRWLLMVNPTRR 296  
Db 351 WSLGVLLYILVSGSLPFDQNLKEURVLRGKYRIPFYMWSTDCENLLKKFLILNPSKGTLL 410



Db 568 TGYLNGGVVDASGIPLPMRYTLP TAASPAPNSCSTSRVG-----RHSLSSSS 619  
Qy 470 P-----EPSESG-----ELLDASDVFSVGDVVEQSPQASGLLLHR---KGI 508  
Db 620 PRSHRPVAISLSDNNPSLANLRCEMMEA-----GGPVGAVGVLASKQLHOTISEFI 674  
Qy 509 LKLNKFSRTALEGTT-----PSTFGSLDQLASHHPAARPSRPS-----GA 549  
Db 675 IKQSTEDCRALLOQSTAVAEKDDPPKAESVGVGPPVPASTPTTSSTAGPSGAPCFGE 734  
Qy 550 VSEDSI---LSSESFD 562  
Db 735 INGTIITKTMSSSSSFD 750

RESULT 8  
T52633  
serine/threonine-specific protein kinase (EC 2.7.1.1-) AKIN11 [validated] - Arabidopsis thaliana  
N:Alternate names: SNF1 protein kinase omolog AKIN11  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 09-Jul-2004  
C:Accession: T52633  
R:Bhalerao, R.P.; Saichert, K.; Bako, L.; Okresz, L.; Szabados, L.; Muranaka, T.; Machid  
Proc. Natl. Acad. Sci. U.S.A. 96, 5322-7, 1999  
A:Title: Regulatory interaction of PR1 WD protein with Arabidopsis SNF1-like protein ki  
A:Reference number: Z25116; MUID:99238528; PMID:10220464  
A:Accession: T52633  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-512 <BHA>  
A:Cross-references: UNIPROT:P92958; EMBL:X99279; PIDN:CAA67671.1  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Gene: AKIN11  
C:Function:  
A:Description: EC 2.7.1.1-; serine/threonine-specific protein kinase AKIN11 [validated, M  
complements SNF1 mutations in yeast  
C:Superfamily: AMP-activated protein kinase; protein kinase homology  
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

Query Match 19.7%; Score 649.5; DB 2; Length 512;  
Best Local Similarity 38.6%; Pred. No. 2.2e-19;  
Matches 152; Conservative 65; Mismatches 126; Indels 51; Gaps 9;

Qy 57 YEFLETGKGYGVKKARE--SSGRLVAIKRKDKIDEQDLHIREIEMSSLNHPH 115  
Db 20 YLKGKTIGISFGVKVKAHVVTHGKVAIKLNRRKIKNMEMEKVREIKILRFMHPH 79  
Qy 116 IIAIHEVFENSSKIVIVMEYASRGDLYDYSERPRLSERDARHFFRQIVSALHYCHQNGI 175  
Db 80 IIRQYEVIEETSDIVVMEYVYKSGELFDYIVKGRLOEDARNFFQIISGVEYCHRMV 139  
Qy 176 VHRDLKLENTLLDANGNIKADFGSLNLYHKGFLOTFCGSPLYASPEIVNGKPYGVPEV 235  
Db 140 VHRDLKPENTLLDANGNIKADFGSLNVMRDGHFLKTCGSPNYPAAPEVSGKLYAGPEV 199  
Qy 236 DSWSLGVLVLYLVHGTMPFDQDHTKLVKQISNGAYREPPK-PSDACGLIRWLLMVPTR 294  
Db 200 DVNSCGVILVALLCGTLPFDDENTPNFKIKGGIYILPSHLSSEARDLIPRLMIVDPVK 259  
Qy 295 RATLEDVASHVWVWVNGYTTGVGEQALREBGGHPSGDFGRASMDWLRSSRPPLLENGAKV 354  
Db 260 RITPEIRQHRWF-----QTHLPRLVAVSPDPDVEQAKKI 294  
Qy 355 CSFFKQHVPGGGSVPGLERQHSUK--KSRKENDMAQNLOQDPAEDTSSRPKSSLLPKP 412  
Db 295 NEETVQEVN-----MGFDRNQVLESRLNRNTQND-----ATVTYILLDNRFRVPS 340  
Qy 413 GILKK--KSSTSSGVEQEDQELRPVDPDTPQPV 444  
Db 341 GYLESEFQETDSG---SNPMR-----TPEAGASPV 368

RESULT 9  
A26030  
serine/threonine-specific protein kinase (EC 2.7.1.1-) SNF1 - yeast (Saccharomyces cerevi  
N:Alternate names: protein YDR477w  
C:Species: Saccharomyces cerevisiae  
C:Date: 20-Aug-1987 #sequence\_revision 20-Aug-1987 #text\_change 09-Jul-2004  
C:Accession: A26030; S69644  
R:Celenza, J.L.; Carlson, M.  
Science 233, 1175-1180, 1986  
A:Title: A yeast gene that is essential for release from glucose repression encodes a p  
A:Reference number: A26030; MUID:86289463; PMID:3526554  
A:Accession: A26030  
A:Molecule type: DNA  
A:Residues: 1-633 <CEL>  
A:Cross-references: UNIPROT:P06782; EMBL:M13971; NID:gi72629; PIDN:AAA35058.1; PID:gi726  
R:Dietrich, F.S.  
submitted to the EMBL Data Library, August 1995  
A:Description: The sequence of S. cerevisiae cosmids 9410, 8035, 8166, and 9787.  
A:Reference number: S69554  
A:Accession: S69644  
A:Molecule type: DNA  
A:Residues: 1-633 <DIE>  
A:Cross-references: EMBL:U33050; NID:g927726; PIDN:AAB64904.1; PID:g927732; GSPDB:GN0000  
C:Genetics:  
A:Gene: SGD:SNF1; MIPS:YDR477w  
A:Cross-references: SGD:S0002885; MIPS:YDR477w  
A:Map position: 4R  
C:Function:  
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin  
A:Note: required for expression of glucose-repressed genes in response to glucose depriv  
C:Superfamily: AMP-activated protein kinase; protein kinase homology  
C:Keywords: ATP; autophosphorylation; magnesium; nucleus; phosphoprotein; phosphotransfe  
E:53-306/Domain: protein kinase homology <KIN>  
E:61-69/Region: protein kinase ATP-binding motif  
F:84,103,177,179/Active site: Lys, Glu, Asp, Lys #status predicted  
F:182,186/Binding site: magnesium (Asn, Asp) #status predicted  
F:210/Binding site: phosphate (Thr) (covalent) (by unidentified kinase) #status predicte

Query Match 19.7%; Score 648.5; DB 1; Length 633;  
Best Local Similarity 33.1%; Pred. No. 2.9e-19;  
Matches 161; Conservative 79; Mismatches 133; Indels 113; Gaps 14;

Qy 47 HHKHNLRH-----RYEFLTLGKTYGVKKKA-RESSGRLV 82  
Db 22 HHHHHHHHGGGNSLNPNKSSLADGAHIGNYQIVKTLGEGSGFKVLAHTTTCQKV 81  
Qy 83 AIKSIKDKIKDEODLL-HIRREIEMSSLNHPHIIAHEVFENSSKIVIVMEYASRGDL 141  
Db 82 ALKIINK-KVLAKSDMQGRIEREISYLLRLRHPHIIKLYDVIKSKOBIIMVIEYAG-NEL 139  
Qy 142 YDYSERPRLSERDARHFFRQIVSALHYCHONGIVHRDLKLENTLLDANGNIKADFGLS 201  
Db 140 FDYIVQDKMSGEARRFFQIISAVEYCHRRKLVHRDLKPENTLLDHLNVLKADFGLS 199  
Qy 202 NLXHGKGFLOTFCGSPLYASPEIVNGKPYGVPEVDSWSLGLVLYLVHGTMPFDQDHTK 261  
Db 200 NIWTDGNTLTCGSPNYPAAPEVSGKLYAGPEVDVWSCGVILVLMCRLLPFPDESIPV 259  
Qy 262 LVQISNGAYREPPKPS-DACGLIRWLLMVPTRATLEDVASHVWVWVNGYTTGVGEQEA 320  
Db 260 LFKNISNGVYTLPKFLSPGAAGLIKRLMIVNPLNRISEIHMQD----- 303  
Qy 321 LREGHPSGDFGRASMDWLRSSRPPLLENGAKVCSFFKQHVPGGGSVPGLERQHSUKK 380  
Db 304 -----DW-----FKVDLP-----EYLLPP 317  
Qy 381 SRKENDMAQNLOQDPAEDTSSRPKSSLLPKGILKKKSSSTSSGVEQEDQELRPVDPDTP 440  
Db 318 DLKPHPEENENDSKDGS---PONDEIDNNLVNLSSTMGVEKDEIVESLESSEDTTP 374  
Qy 441 --GQVPDAVSL--PRKGIKKSKRQRESGYSSPEPSGELLSDASDVFGSDP-----VEQ 493  
Db 375 AFNEIRDVYMLIKENKSLIKDMKANKS-----VSDLDTLFSLQSPPTFQQQS 421

QY 494 KSPQAS 499  
Db 422 KSHQKS 427

## RESULT 10

JC7500  
qik protein - chicken  
N;Alternate names: Qin-induced kinase  
C;Species: Gallus gallus (chicken)  
C;Date: 17-Nov-2000 #sequence\_revision 17-Nov-2000 #text\_change 09-Jul-2004  
C;Accession: JC7500  
R;Xia, Y.; Zhang, Z.; Kruse, U.; Vogt, P.K.; Li, J.  
Biochem. Biophys. Res. Commun. 276, 564-570, 2000  
A;Title: The new serine-threonine kinase, Qik, is a target of the qin oncogene.  
A;Reference number: JC7500  
A;Contents: Embryo fibroblasts  
A;Accession: JC7500  
A;Molecule type: mRNA  
A;Residues: 1-798 <XIA>  
A;Cross-references: UNIPROT:Q91A88; GB:AF219232  
C;Comment: This protein, a member of the AMPK/SNF1 family of serine/threonine kinases,  
A;Gene: qik  
C;Keywords: protein kinase

Query Match 19.7%; Score 648; DB 2; Length 798;  
Best Local Similarity 29.4%; Pred. No. 3.8e-19;  
Matches 195; Conservative 94; Mismatches 223; Indels 152; Gaps 17;

QY 11 SQAPSASALASESARPLADGLIKSPKPLMKQAVKRHHHKNLRRHYEFLETLGKGYGK 70  
Db 5 SEDASVPAPSAQAQPRPLRVGF-----YDIERTLGKGNFAV 39  
QY 71 VKARESSGR-L-VAIKSRDKIKDEQDLLHIREIEIMSSLNHPHIIAHEVFENSCKI 129  
Db 40 VKLARHRTVTKQVAIKIDKTRL-DPSNLEKIREVQIMKLNHPHIIKLVQVNETKML 98  
QY 130 VIVMEYASRGDLYDIYSRPLSRDARHFRQIVSALHYCHQNGIVHRDLKLENILDA 189  
Db 99 YIVTEFAKNGEMFOHDLTNSHGLSEARKEKFWILSAVEYCHSHIHRDLKTENLLDA 158  
QY 190 NGNLIKADFGLSNLYHKGKFTQTCGSPLYASPELVNGKPVVGEVDSWSLGVLLYLIVH 249  
Db 159 NNNIKLADFSGNFYKSGEPLSTWCGSPPYAAPEVTEGKEVTEGPHLDLWSLGVVLYLVC 218  
QY 250 GTMPFFDGDHKTLLVKQISNGAYREPPKPSDAC-GLIRWLLMVNPTRRATLEDVASHWVN 308  
Db 219 GSLPFDGPNLTLRQRLVLEGRFRIPYFNSDECETLIRMLVVDPTKRITISQIKQKWM- 277  
QY 309 WGYTTGVEQEAALREGGHPSGDGRASWADWLRRSSRLP-----LENGAKVCFFQKHVPDG 365  
Db 278 -----QADPSLRQQQSLSFSMQNYSNLGDYNEQVLGI 310  
QY 366 GSTVPLERQHSLLK-----SRKENDMAQLQDPAEDTSSRPGKSSL 408  
Db 311 MQLT-LGIDRQRTVESLQNSVNHFAIYVLLERLKYRSQLSRPRATGQQPRSRSEI 369  
QY 409 ---KLPGILKKKSSGEVQEDQQLRPV-----PDTPGQPV-----PAVSLPLR 452  
Db 370 SNAEMPQDLSLTLRSLILYQQPQSLIQPSLQAECDMDNPLQVFFVPDPNPNGLFR 429  
QY 453 KGILKKQRSGSYSSPEPSEGBLLDASDVFSVGDPEQKSPQASGLLHRKGLKLN 512  
Db 430 NRSISPSSLLETTI---SSEVRQKLED---EIKAYDHPIRPS-----N 469  
QY 513 KFSFTALEGTPSTFGSLDOLASHPAARPSRPSGAVSEDSILSESFDQ----- 563  
Db 470 TSSRHLEAVTHFYQHAPPCIVISSAS-----PTEGTSSSCLTSSNDSVALSSCLA 525  
QY 564 -----LDLPERLPETPLRGCVSVNLT-----RGLQPPSPSEGLKR 597

Db 526 QGVMTGSPATARTWTSAPLASQSDAPVLQVQGWGASLLPVSFQEGRRASDTSLTQGLKA 585  
QY 598 WWOQE 601  
Db 586 FRQQ 589

## RESULT 11

T10449  
Probable serine/threonine-specific protein kinase (EC 2.7.1.1) - cucumber  
N;Alternate names: SNF1-related protein kinase  
C;Species: Cucumis sativus (cucumber)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C;Accession: T10449  
R;Gumpel, N.J.  
submitted to the EMBL Data Library, December 1996  
A;Reference number: Z17020  
A;Accession: T10449  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-504 <GUM>  
A;Cross-references: UNIPROT:P93113; EMBL:Y10036  
A;Experimental source: cv. Masterpiece; coryledon  
C;Function:  
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine  
C;Superfamily: AMP-activated protein kinase; protein kinase homology  
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
F;6-260/Domain: protein kinase homology <KIN>

Query Match 19.5%; Score 641; DB 2; Length 504;  
Best Local Similarity 50.4%; Pred. No. 4.7e-19;  
Matches 127; Conservative 47; Mismatches 76; Indels 2; Gaps 2;

QY 57 YEFLETLGKTYGKVKKARES-SGRLVATKSIKDKIKDQDQLLHIREIEIMSSLNHPH 115  
Db 8 YKLGKTLGIGSGFKVIAEHALTGKVAIKILNRRKIKNDMEKVRREIKILRLFWHPH 67  
QY 116 IIAIHEVFENSSKIVIMVEYASRGDLYDIYSRPLSRDARHFRQIVSALHYCHQNGI 175  
Db 68 IIRLYEVIETPSDIYVMEYVKSGLDFDIYVEKGRQLQEDARNFFQIISGVEYCHRMV 127  
QY 176 VHRDLKLENILLDANGNIKIAFDGLSNLYHKGFLOFTFCGSPLYASPEIYNGKPYGPEV 235  
Db 128 VHRDLKLENILLDSKCNVKIADFGLSNIMRDGHFLKTCGSPNYAAPEVTSGLKYAGPEV 187  
QY 236 DWSLGLVLLYLIVHGTMPDPGQDHKTLLVKQISNGAYREPPK-PSDACGLIRWLLMVNPT 294  
Db 188 DVWSCGVLLYALLCGLTLPDDENIPNLFKKIGGIYTLPSHLSSGARELIPSLMLVVDPMK 247  
QY 295 RATLEDVASHWW 306  
Db 248 RIITPEIRQHPW 259

## RESULT 12

JC1446  
serine/threonine-specific protein kinase (EC 2.7.1.1) AK21 - Arabidopsis thaliana  
N;Alternate names: protein kinase SNF1 homolog  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C;Accession: JC1446; S58266; S66334  
R;LeGuen, L.; Thomas, M.; Bianchi, M.; Halford, N.G.; Kreis, M.  
Gene 120, 249-254, 1992  
A;Title: Structure and expression of a gene from Arabidopsis thaliana encoding a protein  
A;Reference number: JC1446; MUID:93013041; PMID:1339373  
A;Accession: JC1446  
A;Molecule type: DNA  
A;Residues: 1-512 <LEG>  
A;Cross-references: UNIPROT:Q38997; GB:M93023; NID:g166599; PIDN:AAA32736.1; PID:g166600  
R;Thummler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.  
submitted to the EMBL Data Library, May 1995  
A;Description: differential accumulation of the transcripts of 22 novel protein kinase g  
A;Reference number: S58256



C;Superfamily: AMP-activated protein kinase; protein kinase homology  
C;Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase  
F;14-269/Domain: protein kinase homology <KIN>  
F;22-30/Region: protein kinase ATP-binding motif  
F;45,64,139,141/Active site: Lys, Glu, Asp, Lys #status predicted  
F;144,148/Binding site: magnesium (Asn, Asp) #status predicted

```
Query Match      19.1%; Score 630.5; DB 1; Length 552;
Best Local Similarity 31.5%; Pred. No. 1.3e-18;
Matches 176; Conservative 91; Mismatches 199; Indels 93; Gaps 17;

Qy 45 KRHHKHNLRHRYFLETGLGTYGKVK-KARESSGRLVAIKSRKDKIKDEQDQLLHRR 103
Db 5 QKHGRVKIGH-VYLGDTLGVGTGKVGKEGHQLTGHKVAVKILNRQKIRSLDVVGKIKR 63

Qy 104 EIEIMSSLNHPHIIAIEHVFENSSKIIVIMYASRGDLYIISRPRLSERDARHFRQI 163
Db 64 EIQNLKLFRRPHIILKYQISTPTDFVMVEYVSGGELFDYICKHGRVEEMEARRLFQOI 123

Qy 164 VSALHYCHONGIVHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSPLYASPE 223
Db 124 LSAVDYCHRMVVRDLKPNVLDADNNAKIADFGLSNMMSGDEFLRTSCGSPNYAAPE 183

Qy 224 IVNGKPYVGPVDSWSLGLVLLYILVHGTMPDGDHKTIVKQISNGAYREPPKPS-DACG 282
Db 184 VISGRLYAGPEVDIWSGVLVYALLCGTLFPDDEHVPTLFKKIRGGVFIPEYLNRSVAT 243

Qy 283 LIIRLLMVPNTRATLEDVASHWVWV-----GYTTGVGEQEARLE----- 323
Db 244 LLMHMLQVDPKLRATIKDIREHEWFKQDLPSPYLPEDPSYDANVIDDEAVKEVERPECT 303

Qy 324 -----GHPGSGDFGRASMDLRRSSRPLENGAKVCSFFKQHPVGGSTV----- 369
Db 304 ESEVWNSLYSGDPQ---DQLAVAVHLIIDNRINWQASE---FYLASSPPSGSFMDDSAM 357

Qy 370 ---PGL-----ERQSL-----KSKRKENDMAQNLQGDPAEDTSSRPKSKSLKLPKGIKXK 418
Db 358 HIPGLKPHPERMPPLIADSPKARCLDAL-----NTTKPKSLAVKAKWHLGIR 407

Qy 419 SSTSSGSEVEDPQLRPVPTPGOPVPAVSLLPK-----GILKSKRQ-----ESGYISS 469
Db 408 SQSKPYDIMEAIVRYAMKQLDFEWKVVNAYHLRVRKKNPVVTGNYVKMSLQLYLVNRSYLL 467

Qy 470 PESESGELLDASDVFGSDPVEQKSPQASGLLHRRKILKNGKFGRTALE----- 521
Db 468 DFKSIDDEVVEQR-----SGSSTPQSCSAAG--LHRP-----RSSPSTTAESHLSGSL 516

Qy 522 --GTPPTFGSLDQLASH 538
Db 517 TGSITGSLSSVSPRLGSH 535
```

## RESULT 15

S72513  
FOG2 protein - yeast (*Kluyveromyces marxianus* var. *lactis*)  
C;Species: *Kluyveromyces marxianus* var. *lactis*, Candida spheerica  
C;Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 09-Jul-2004  
C;Accession: S72513  
R;Goffin, P.; Picarelli, A.; Donnini, C.; Lodi, T.; Puglisi, P.P.; Ferrero, I.  
Curr. Genet. 29, 316-326, 1996  
A;Title: FOG1 and FOG2 genes, required for the transcriptional activation of glucose-rep  
A;Reference number: S72513; MUID:96171514; PMID:8598052  
A;Accession: S72513  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-602 <GOF>  
A;Cross-references: UNIPROT:P87209  
A;Note: the source is designated as *Kluyveromyces lactis*  
C;Genetics:  
A;Gene: FOG2  
C;Function:  
A;Description: probably involved in the regulation of glucose-repressible gene expressio  
C;Superfamily: AMP-activated protein kinase; protein kinase homology

F;33-286/Domain: protein kinase homology <KIN>

```
Query Match      19.1%; Score 627.5; DB 2; Length 602;
Best Local Similarity 35.6%; Pred. No. 1.9e-18;
Matches 155; Conservative 68; Mismatches 125; Indels 87; Gaps 14;

Qy 47 HHHKHNLRH-----RYELETGLGTYGKVKKARE-SSGRLVAIKSRKDKIKDSODL 98
Db 18 HHORQLTNHAQGHGKIYQIKITLGEFGKVLAVHIHSTGQKVALKIINK-KVLAKSDM 76

Qy 99 L-HIRREIETMSSLNHPHIIAIEHVFENSSKIIVIMYASRGDLYIISRPRLSERDAR 157
Db 77 QGRIEREISVLRLLRHPHIIKLYDVITKSQDSIIIMVIEYAG-NELFDYIVQDKMPEQEAR 135

Qy 158 HFRQIVSALHYCHONGIVHRDLKLENILLDANGNIKIADFGLSNLYHKGFLOTFCGSP 217
Db 136 RFFQQIISAVDYCHRHKIVHRDLKPNVLDADNNAKIADFGLSNIMTGDNFLTSCGSP 195

Qy 218 LYASPEIVNGKPYVGPVDSWSLGLVLLYILVHGTMPDGDHKTIVKQISNGAYREPPKP 277
Db 196 NYAAPEVISGKLYAGPEVDVWSSGVLYVLMCRRLPDDDESIPVLFKNISNGVYTIENFL 255

Qy 278 SD-ACGLIRLLMVPNTRATLEDVASHWVWVNNWGYTTGVGEQEARLEGHPGSGDFGRAS 336
Db 256 SQGAASLIKKQLIVNPNRITVHEIMQDEWF-----KVDL 290

Qy 337 ADMLRRSSRPLENGAKVCSFFKQHPVGGSTVP-----G 371
Db 291 PDYLVPAESTHQENS-----ESKTEDEGSPVPLELIDSLVQTLNMTGYDVDEIYEA 343

Qy 372 LERQH---SLKSKR-----KE-----NDMAQLQGDPAEDT---SSRPG---KSSLKLPKG 413
Db 344 LESEDEHPSLNEIRDAYQLIKENLNINDIKVKKQSQNSLDLTFLSQSPTFEQSLHAPPG 403

Qy 414 ILKKKSTSSGSEVQE 428
Db 404 SKNRHSHRSKRTOQ 418
```

Search completed: January 22, 2005, 06:43:51  
Job time : 42.5452 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2005, 06:43:09 ; Search time 146.329 Seconds  
(without alignments)  
1557.950 Million cell updates/sec

Title: US-09-980-464-11  
Perfect score: 3293  
Sequence: 1 MESVALLQRPQAPSASALA.....DCQEVTAAYRQALGICKLS 631

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3293	100.0	631	14	US-10-355-975-11
2	3186	96.8	640	16	US-10-322-281-23
3	3068.5	93.2	630	16	US-10-343-514-41
4	2824.5	85.8	628	9	US-09-963-159-2
5	2824.5	85.8	628	15	US-10-423-543-44
6	2824.5	85.8	628	17	US-10-370-715B-640
7	2824.5	85.8	672	17	US-10-618-941-77
8	2720.5	82.6	594	15	US-10-311-034-12
9	2712.5	82.4	616	16	US-10-322-281-26
10	1781.5	54.1	406	16	US-10-363-829-404
11	1781.5	54.1	406	16	US-10-363-829-504
12	1699.5	51.6	661	9	US-09-780-949-2
13	1699.5	51.6	661	9	US-09-780-949-6

14	1699.5	51.6	661	14	US-10-354-358-82	Sequence 82, Appl
15	1699.5	51.6	661	17	US-10-723-860-2553	Sequence 2553, Ap
16	1325	40.2	251	16	US-10-343-514-103	Sequence 103, App
17	1301.5	39.5	530	9	US-09-836-392-20	Sequence 20, Appl
18	1296	39.4	251	16	US-10-343-514-50	Sequence 50, Appl
19	1276	38.7	251	16	US-10-343-514-101	Sequence 101, App
20	1111.5	33.8	252	16	US-10-343-514-102	Sequence 102, App
21	785.5	23.9	246	10	US-09-898-837A-29	Sequence 29, Appl
22	748	22.7	703	15	US-10-016-248-70	Sequence 70, Appl
23	748	22.7	752	9	US-09-835-081-2	Sequence 2, Appli
24	748	22.7	752	15	US-10-258-106-16	Sequence 16, Appl
25	748	22.7	752	15	US-10-276-645-7	Sequence 7, Appli
26	748	22.7	825	15	US-10-425-114-54516	Sequence 54516, A
27	744	22.6	639	15	US-10-016-248-71	Sequence 71, Appl
28	744	22.6	639	15	US-10-016-248-72	Sequence 72, Appl
29	744	22.6	688	15	US-10-276-645-8	Sequence 8, Appli
30	744	22.6	752	17	US-10-618-941-79	Sequence 79, Appl
31	743.5	22.6	769	15	US-10-363-616-403	Sequence 403, App
32	739	22.4	639	15	US-10-016-248-73	Sequence 73, Appl
33	739	22.4	688	14	US-10-161-565-28	Sequence 28, Appl
34	739	22.4	688	14	US-10-161-565-29	Sequence 29, Appl
35	735	22.4	724	15	US-10-276-645-5	Sequence 5, Appli
36	732.5	22.2	660	15	US-10-276-645-6	Sequence 6, Appli
37	703.5	21.4	729	14	US-10-142-356-11	Sequence 11, Appl
38	703.5	21.4	729	14	US-10-195-101-33	Sequence 33, Appl
39	703.5	21.4	729	14	US-10-161-565-26	Sequence 26, Appl
40	701	21.3	140	16	US-10-343-514-47	Sequence 47, Appl
41	701	21.3	744	9	US-09-835-081-4	Sequence 4, Appli
42	700.5	21.3	1038	15	US-10-424-599-274878	Sequence 274878, A
43	698.5	21.2	624	15	US-10-425-114-70100	Sequence 70100, A
44	698.5	21.2	744	9	US-09-919-585-3	Sequence 3, Appli
45	698	21.2	698	15	US-10-016-248-22	Sequence 22, Appl

#### ALIGNMENTS

#### RESULT 1

US-10-355-975-11  
; Sequence 11, Application US/10355975  
; Publication No. US20030162277A1  
; GENERAL INFORMATION:  
; APPLICANT: Immunex Corporation  
; APPLICANT: Bird, Timothy A.  
; APPLICANT: Virca, G. Duke  
; APPLICANT: Martin, Unja  
; APPLICANT: Anderson, Dirk M.  
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES  
; FILE REFERENCE: 2923-A  
; CURRENT FILING DATE: 2003-01-30  
; PRIOR FILING DATE: 2003-01-30  
; PRIOR APPLICATION NUMBER: US/09/579,664B  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 11  
; LENGTH: 631  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; US-10-355-975-11

Query Match 100.0%; Score 3293; DB 14; Length 631;  
Best Local Similarity 100.0%; Pred No. 1.4e-206;  
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy . . . 1 MESVALLQRPQAPSASALASARPLADGLIKSPKPLMKQAVKRRHHKHLRHYEFL 60

Db . . . 1 MESVALLQRPQAPSASALASARPLADGLIKSPKPLMKQAVKRRHHKHLRHYEFL 60

Qy . . . 61 ETLLGKGTGYGVKKARESSGLRLVAIKRKDKIKDEQDLLHIREIEMSSLNPHIATIH 120

Db . . . 61 ETLLGKGTGYGVKKARESSGLRLVAIKRKDKIKDEQDLLHIREIEMSSLNPHIATIH 120



Qy 121 EVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIVHRDL 180  
Db 121 EVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIVHRDL 180  
Qy 181 KLENILLDANGNIKIADFGLSNLHYHKGKFLQTFGSGPLYASPEIVNGKPYVGPPEVDSWSL 240  
Db 181 KLENILLDANGNIKIADFGLSNLHYHKGKFLQTFGSGPLYASPEIVNGKPYVGPPEVDSWSL 240  
Qy 241 GVLILYLVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDAGCLIRWLLMVPNTRATLED 300  
Db 241 GVLILYLVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDAGCLIRWLLMVPNTRATLED 300  
Qy 301 VASHWWNWGTYTGVGEQALREGGHPGSDPGRASMDWLRSSRPILLENGAKVCSPFKQ 360  
Db 301 VASHWWNWGTYTGVGEQALREGGHPGSDPGRASMDWLRSSRPILLENGAKVCSPFKQ 360  
Qy 361 HVPGGGTVPGLEQRHSLKSKRKNENDMAQNLOQDPAEDTSRRPGKSKLPLKGLKKSS 420  
Db 361 HVPGGGTVPGLEQRHSLKSKRKNENDMAQNLOQDPAEDTSRRPGKSKLPLKGLKKSS 420  
Qy 421 TSSEVOEDPQELRPVDPDTPGQVPVAVSLLPRKGLKKSRQESGYSSPEPSESGELLD 480  
Db 421 TSSEVOEDPQELRPVDPDTPGQVPVAVSLLPRKGLKKSRQESGYSSPEPSESGELLD 480  
Qy 481 ASDVFVSGDPVEQKSPQASGLLLHRRKGLKNGKFSRTALEGTPSTFGSLDOLASSHPA 540  
Db 481 ASDVFVSGDPVEQKSPQASGLLLHRRKGLKNGKFSRTALEGTPSTFGSLDOLASSHPA 540  
Qy 541 ARPSRPGAVSEDSILSESFDQLDLPERLPETPLRGCVSDNLRGLQEPPESEGLKRWQ 600  
Db 541 ARPSRPGAVSEDSILSESFDQLDLPERLPETPLRGCVSDNLRGLQEPPESEGLKRWQ 600  
Qy 601 ESLGDSFSLTDCQEVTAAYRQALGICSKLS 631  
Db 601 ESLGDSFSLTDCQEVTAAYRQALGICSKLS 631

## RESULT 2

US-10-322-281-23  
; Sequence 23, Application US/10322281  
; Publication No. US20040126762A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: 529452001000  
; CURRENT APPLICATION NUMBER: US/10/322,281  
; CURRENT FILING DATE: 2002-12-17  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 640  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-322-281-23

Query Match 96.8%; Score 3186; DB 16; Length 640;  
Best Local Similarity 98.3%; Pred. No. 1.4e-199;  
Matches 620; Conservative 0; Mismatches 1; Indels 10; Gaps 5;

Qy 1 MESVALLQRPQAPSALASESARPLADGLIKSPKLMKKQAVKRHHKHNLRHYEFL 60  
Db 20 MESV--LQRPQAPSAS--ASESARPLADGLIKSPKLMKKQAVKRHHKHNLRHYEFL 75  
Qy 61 ETLGKTYGKVKKARESSGRVVAIKSRDKIKDEQDLHRRREIEMSSLNHPHIIAIIH 120  
Db 76 ETLGKTYGKVKKARESSGRVVAIKSRDKIKDEQDLHRRREIEMSSLNHPHIIAIIH 135  
Qy 121 EVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIVHRDL 180  
Db 136 EVFENSSKIVIMEYASRGDLVDYIS--PRLS--DARHFFRQIVSALHYCHONGIVHRDL 191  
Qy 181 KLENILLDANGNIKIADFGLSNLHYHKGKFLQTFGSGPLYASPEIVNGKPYVGPPEVDSWSL 240

Db 192 KLENILLDANGNIKIADFGLSNLHYHKGKFLQTFGSGPLY--DEIVNGKPYVGPPEVDSWSL 249  
Qy 241 GVLILYLVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDAGCLIRWLLMVPNTRATLED 300  
Db 250 GVLILYLVHGTMPFDGQDHKTLVKQISNGAYREPPKPSDAGCLIRWLLMVPNTRATLED 309  
Qy 301 VASHWWNWGTYTGVGEQALREGGHPGSDPGRASMDWLRSSRPILLENGAKVCSPFKQ 360  
Db 310 VASHWWNWGTYTGVGEQALREGGHPGSDPGRASMDWLRSSRPILLENGAKVCSPFKQ 369  
Qy 361 HVPGGGTVPGLEQRHSLKSKRKNENDMAQNLOQDPAEDTSRRPGKSKLPLKGLKKSS 420  
Db 370 HVPGGGTVPGLEQRHSLKSKRKNENDMAQNLOQDPAEDTSRRPGKSKLPLKGLKKSS 429  
Qy 421 TSSEVOEDPQELRPVDPDTPGQVPVAVSLLPRKGLKKSRQESGYSSPEPSESGELLD 480  
Db 430 TSSEVOEDPQELRPVDPDTPGQVPVAVSLLPRKGLKKSRQESGYSSPEPSESGELLD 489  
Qy 481 ASDVFVSGDPVEQKSPQASGLLLHRRKGLKNGKFSRTALEGTPSTFGSLDOLASSHPA 540  
Db 490 ASDVFVSGDPVEQKSPQASGLLLHRRKGLKNGKFSRTALEGTPSTFGSLDOLASSHPA 549  
Qy 541 ARPSRPGAVSEDSILSESFDQLDLPERLPETPLRGCVSDNLRGLQEPPESEGLKRWQ 600  
Db 550 ARPSRPGAVSEDSILSESFDQLDLPERLPETPLRGCVSDNLRGLQEPPESEGLKRWQ 609  
Qy 601 ESLGDSFSLTDCQEVTAAYRQALGICSKLS 631  
Db 610 ESLGDSFSLTDCQEVTAAYRQALGICSKLS 640

## RESULT 3

US-10-343-514-41  
; Sequence 41, Application US/10343514  
; Publication No. US20040132025A1  
; GENERAL INFORMATION:

; APPLICANT: DRUCKER, Daniel J.  
; APPLICANT: ROSEN, Cheryl F.  
; APPLICANT: LEBEVRE, Diana L.  
; TITLE OF INVENTION: AMPK-RELATED SERINE/THREONINE KINASE, DESIGNATED SNARK  
; FILE REFERENCE: DPA-DRUC2/PCT  
; CURRENT APPLICATION NUMBER: US/10/343,514  
; CURRENT FILING DATE: 2003-01-31  
; PRIOR APPLICATION NUMBER: PCT/CA01/01109  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR APPLICATION NUMBER: US 60/222,650  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/274,613  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: CA 2,340,780  
; PRIOR FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 109  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 41  
; LENGTH: 630  
; TYPE: PRT  
; ORGANISM: RAT  
US-10-343-514-41

Query Match 93.2%; Score 3068.5; DB 16; Length 630;  
Best Local Similarity 94.0%; Pred. No. 6.4e-192;  
Matches 593; Conservative 12; Mismatches 25; Indels 1; Gaps 1;

Qy 1 MESVALLQRPQAPSALASESARPLADGLIKSPKLMKKQAVKRHHKHNLRHYEFL 60  
Db 1 MESVALLHRRGNLAPSALATESAREPLADRLIKSPKLMKKQAVKRHHKHNLRHYEFL 60  
Qy 61 ETLGKTYGKVKKARESSGRVVAIKSRDKIKDEQDLHRRREIEMSSLNHPHIIAIIH 120  
Db 61 ETLGKTYGKVKKARESSGRVVAIKSRDKIKDEQDLHRRREIEMSSLNHPHIIAIIH 120  
Qy 121 EVFENSSKIVIMEYASRGDLVDYISERPRLSERDARHFFRQIVSALHYCHONGIVHRDL 180

Db 121 EVFNSKIVMEYASRGDLYISERPRNERDARHFRQIVSALHYCHONGIVHRDL 180  
Qy 181 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPVDSWSL 240  
Db 181 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPVDSWSL 240  
Qy 241 GVLLYILVHGTMPDGDHKTIVKQISNGAYRPPKPSDACGLIRWLLMNPTRATLED 300  
Db 241 GVLLYILVHGTMPDGDHKTIVKQISNGAYRPPKPSDACGLIRWLLMNPTRATLED 300  
Qy 301 VASHWVNWGTYTTCVGEQALREGGHPGSDGRASMDLRRSRPILLENGAKVCSFPKQ 360  
Db 301 VASHWVNWGTYTTCVGEQALREGGHPGSDGRASMDLRRSRPILLENGAKVCSFPKQ 360  
Qy 361 HVPGGSTVPLERQHSLSKSRKENDMAQNLQDPAEDTSSRPCKSKLPLKGLKKKSS 420  
Db 361 HVPGGSTVPLERQHSLSKSRKENDMAQNLQDPAEDTSSRPCKSKLPLKGLKKKSS 420  
Qy 421 TSSEVOEDQELRPVDPDTPGQPVPAVSLPRKGLKKSRQREGSGYSSPEPSGELL 480  
Db 421 PSSGEVQEGQELRPVDPDTPGQPVPAVSLPRKGLKKSRQREGSGYSSPEPSGELL 480  
Qy 481 ASDVFGSDPVEQKSPQASGLLHRRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540  
Db 481 AGDVFGSDPVEQKSPQASGLLHRRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540  
Qy 541 ARPSRPSGAVSEDSILSSESFQDLDLPERLPETPLRGCVSDNLRGLQPPSEGLKRWQ 600  
Db 541 ARPSRPSGAVSEDSILSSESFQDLDLPERLPETPLRGCVSDNLRGLQPPSEGLKRWQ 600  
Qy 601 ESLGDSRFSLTDCQEVTAAYRQALGICSKLS 631  
Db 600 ESLGDSRFSLTDCQEVTAAYRQALGICSKLS 630

## RESULT 4

US-09-963-159-2  
; Sequence 2, Application US/09963159  
; Patent No. US20020077312A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Galvin, Katherine M.  
; TITLE OF INVENTION: 3700, A NOVEL HUMAN PROTEIN KINASE AND USES THEREFOR  
; FILE REFERENCE: 10147-50U1  
; CURRENT APPLICATION NUMBER: US/09/963,159  
; CURRENT FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 60/234,922  
; PRIOR FILING DATE: 2000-09-25  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 628  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-963-159-2

Query Match 85.8%; Score 2824.5; DB 9; Length 628;  
Best Local Similarity 85.7%; Pred. No. 5.4e-176;  
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;  
Qy 1 MESVALLQRFPSQAPASALASESARPLADGLIKSPKMKKQAVKRHHKHLNLRHYEFL 60  
Db 1 MESLVFARRSGPTPS----AAELARPLAEGLIKSPKMKKQAVKRHHKHLNLRHYEFL 56  
Qy 61 ETLGKGYGVKKARESSGRBLVAIKSTRKDKIDEQDLHRIEIMSSLNPHIATIH 120  
Db 57 ETLGKGYGVKKARESSGRBLVAIKSTRKDKIDEQDLHRIEIMSSLNPHIATIH 116  
Qy 121 EVFNSKIVMEYASRGDLYISERPRLSERDARHFRQIVSALHYCHONGIVHRDL 180  
Db 117 EVFNSKIVMEYASRGDLYISERPRLSERDARHFRQIVSALHYCHONGIVHRDL 176

Qy 181 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYVGPVDSWSL 240  
Db 177 KLENILLDANGNIKIADFGLSNLYHKGKFLQTCGSPLYASPEIVNGKPYTGPVDSWSL 236  
Qy 241 GVLLYILVHGTMPDGDHKTIVKQISNGAYRPPKPSDACGLIRWLLMNPTRATLED 300  
Db 237 GVLLYILVHGTMPDGDHKTIVKQISNGAYRPPKPSDACGLIRWLLMNPTRATLED 296  
Qy 301 VASHWVNWGTYTTCVGEQALREGGHPGSDGRASMDLRRSRPILLENGAKVCSFPKQ 360  
Db 297 VASHWVNWGTYTTCVGEQALREGGHPGSDGRASMDLRRSRPILLENGAKVCSFPKQ 356  
Qy 361 HVPGGSTVPLERQHSLSKSRKENDMAQNLQDPAEDTSSRPCKSKLPLKGLKKKSS 420  
Db 357 HAPGGGTTPLERQHSLSKSRKENDMAQNLQDPAEDTSSRPCKSKLPLKGLKKKSS 416  
Qy 421 TSSEVOEDQELRPVDPDTPGQPVPAVSLPRKGLKKSRQREGSGYSSPEPSGELL 480  
Db 417 ASAGVQEDQELRPVDPDTPGQPVPAVSLPRKGLKKSRQREGSGYSSPEPSGELL 473  
Qy 481 ASDVFGSDPVEQKSPQASGLLHRRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540  
Db 474 AGDVFGSDPVEQKSPQASGLLHRRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 533  
Qy 541 ARPSRPSGAVSEDSILSSESFQDLDLPERLPETPLRGCVSDNLRGLQPPSEGLKRWQ 596  
Db 534 ARASRPSGAVSEDSILSSESFQDLDLPERLPETPLRGCVSDNLRGLQPPSEGLKRWQ 593  
Qy 597 RHWQESIGDSFSLTDCQEVTAAYRQALGICSKLS 631  
Db 594 RWRQDPLGDSFSLTDCQEVTAAYRQALGICSKLS 628

## RESULT 5

US-10-423-543-44  
; Sequence 44, Application US/10423543  
; Publication No. US2004005835A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Libermann, Rosana K.  
; APPLICANT: Hunter, John J.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Rudolph-Owen, Laura A.  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Tsai, Fong-Ying  
; APPLICANT: Galvin, Katherine M.  
; APPLICANT: Chun, Miyoung  
; APPLICANT: Williamson, Mark J.  
; APPLICANT: Bandar, Rajasekhar  
; TITLE OF INVENTION: NOVEL 21910, 56634, 55053, 2504, 15977,  
; TITLE OF INVENTION: 14760, 25501, 17903, 3700, 21529, 26176, 26343, 56638,  
; TITLE OF INVENTION: 18610, 33217, 21967, h1983, m1983, 38555 OR 593 MOLECULES  
; FILE REFERENCE: MPI03-0230NMIM  
; CURRENT APPLICATION NUMBER: US/10/423,543  
; CURRENT FILING DATE: 2003-04-25  
; PRIOR APPLICATION NUMBER: US 10/278,036  
; PRIOR FILING DATE: 2002-10-22  
; PRIOR APPLICATION NUMBER: US 09/711,216  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: US 60/205,447  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 10/012,055  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/248,325  
; PRIOR FILING DATE: 2000-11-14  
; PRIOR APPLICATION NUMBER: US 10/003,690  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/248,893  
; PRIOR FILING DATE: 2000-11-15  
; PRIOR APPLICATION NUMBER: US 09/797,039

; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/186,061  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 10/217,168  
; PRIOR FILING DATE: 2002-08-12  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44  
; LENGTH: 628  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-423-543-44

Query Match 85.8%; Score 2824.5; DB 15; Length 628;  
Best Local Similarity 85.7%; Pred. No. 5.4e-176;  
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;  
Qy 1 MESVALLQRPQAPSASALASESARPLADGLIKSPKPLMKQAVKRRHHKHNLRHYEFL 60  
Db 1 MESLVFARRSGPTPS----AAELARPLAELGKSPKPLMKQAVKRRHHKHNLRHYEFL 56  
Qy 61 ETLGKGTGKVKKARESSGRGLVAIKSRKDKIKDEQDLHRRREIEMSSLNHPHIIAIIH 120  
Db 57 ETLGKGTGKVKKARESSGRGLVAIKSRKDKIKDEQDLHRRREIEMSSLNHPHIIAIIH 116  
Qy 121 EVFENSISKIVIMVEYASRGDLVDYISERPLSRDARHFRQIVSALHYCHONGIVHRDL 180  
Db 117 EVFENSISKIVIMVEYASRGDLVDYISERQQLSREARHFRQIVSAVHYCHONRVHRDL 176  
Qy 181 KLENILLDANGNIKIADFGLSNLYHGKFLQTCGSPLYASPELVNGKPYVGPVEVDSWSL 240  
Db 177 KLENILLDANGNIKIADFGLSNLYHGKFLQTCGSPLYASPELVNGKPYTGPVEVDSWSL 236  
Qy 241 GVLLYILVHGTMPPDGDHKTILVKQISNGAYREPPKPSDACGLIRWLLMVPTRATLED 300  
Db 237 GVLLYILVHGTMPPDGDHKTILVKQISNGAYREPPKPSDACGLIRWLLMVPTRATLED 296  
Qy 301 VASHWWNVGTYTGVGEQEARREGHSPGDPGRASMDWLRSSRPLLENGAKVCSFFKQ 360  
Db 297 VASHWWNVGTYATRVGEQEARREGHSPGDSARASMDWLRSSRPLLENGAKVCSFFKQ 356  
Qy 361 HVPGGSTVPCLEHRSKSKRKNENDMAQNLQGPDAEDTSRRPKSSIKLPKGLKKKVS 420  
Db 357 HAPGGSTTTPGLEHRSKSKRKNENDMAQSLHSSTADDTAHRPKSNLKLKPKGLKKKVS 416  
Qy 421 TSSEGEQEDPQELRPVDPDPQPPAVSVLLPRKGLKSKRQRESGYSSPEPSESGLLD 480  
Db 417 ASAEVQEDPPELSPIPASPGQAAP---LLPKGILKKPRQRESGYSSPEPSESGLLD 473  
Qy 481 ASDVFVSGDPVEQKSPQASGLLLHRRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540  
Db 474 AGDVFSVSGDPKEQKPPQASGLLLHRRKGLKNGKFSQTALEAAPTTFGSLDELAPRPL 533  
Qy 541 ARPSRPSGAVSEDSILSESFDQLDLPERLPETPLRGCVSDNLRGLEQPPSEG---LK 596  
Db 534 ARASRPSGAVSEDSILSESFDQLDLPERLPETPLRGCVSDNLTGLEEPPSGGCLR 593  
Qy 597 RWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631  
Db 594 RWRQDPLGDSFSLTDCQEVTAAYRQALRVCSKLT 628

## RESULT 6

US-10-370-715B-640  
; Sequence 640, Application US/10370715B  
; Publication No. US20040258678A1  
; GENERAL INFORMATION:  
; Patin Docket Preview  
; APPLICANT: BODARY, SARAH C.  
; APPLICANT: CLARK, HILLARY  
; APPLICANT: BRISDELL, HUNTE  
; APPLICANT: JACKMAN, JANET

; APPLICANT: SCHOENFELD, JILL R.  
; APPLICANT: WILLIAMS, P. MICKEY  
; APPLICANT: WOOD, WILLIAM I.  
; APPLICANT: WU, THOMAS D.  
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune  
; FILE REFERENCE: PI948R1-US  
; CURRENT APPLICATION NUMBER: US/10/370,715B  
; NUMBER OF SEQ ID NOS: 742  
; NUMBER OF SEQ ID NOS: 742  
; SEQ ID NO 640  
; LENGTH: 628  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-370-715B-640

Query Match 85.8%; Score 2824.5; DB 17; Length 628;  
Best Local Similarity 85.7%; Pred. No. 5.4e-176;  
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;  
Qy 1 MESVALLQRPQAPSASALASESARPLADGLIKSPKPLMKQAVKRRHHKHNLRHYEFL 60  
Db 1 MESLVFARRSGPTPS----AAELARPLAELGKSPKPLMKQAVKRRHHKHNLRHYEFL 56  
Qy 61 ETLGKGTGKVKKARESSGRGLVAIKSRKDKIKDEQDLHRRREIEMSSLNHPHIIAIIH 120  
Db 57 ETLGKGTGKVKKARESSGRGLVAIKSRKDKIKDEQDLHRRREIEMSSLNHPHIIAIIH 116  
Qy 121 EVFENSISKIVIMVEYASRGDLVDYISERPLSRDARHFRQIVSALHYCHONGIVHRDL 180  
Db 117 EVFENSISKIVIMVEYASRGDLVDYISERQQLSREARHFRQIVSAVHYCHONRVHRDL 176  
Qy 181 KLENILLDANGNIKIADFGLSNLYHGKFLQTCGSPLYASPELVNGKPYVGPVEVDSWSL 240  
Db 177 KLENILLDANGNIKIADFGLSNLYHGKFLQTCGSPLYASPELVNGKPYTGPVEVDSWSL 236  
Qy 241 GVLLYILVHGTMPPDGDHKTILVKQISNGAYREPPKPSDACGLIRWLLMVPTRATLED 300  
Db 237 GVLLYILVHGTMPPDGDHKTILVKQISNGAYREPPKPSDACGLIRWLLMVPTRATLED 296  
Qy 301 VASHWWNVGTYTGVGEQEARREGHSPGDPGRASMDWLRSSRPLLENGAKVCSFFKQ 360  
Db 297 VASHWWNVGTYATRVGEQEARREGHSPGDSARASMDWLRSSRPLLENGAKVCSFFKQ 356  
Qy 361 HVPGGSTVPCLEHRSKSKRKNENDMAQNLQGPDAEDTSRRPKSSIKLPKGLKKKVS 420  
Db 357 HAPGGSTTTPGLEHRSKSKRKNENDMAQSLHSSTADDTAHRPKSNLKLKPKGLKKKVS 416  
Qy 421 TSSEGEQEDPQELRPVDPDPQPPAVSVLLPRKGLKSKRQRESGYSSPEPSESGLLD 480  
Db 417 ASAEVQEDPPELSPIPASPGQAAP---LLPKGILKKPRQRESGYSSPEPSESGLLD 473  
Qy 481 ASDVFVSGDPVEQKSPQASGLLLHRRKGLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540  
Db 474 AGDVFSVSGDPKEQKPPQASGLLLHRRKGLKNGKFSQTALEAAPTTFGSLDELAPRPL 533  
Qy 541 ARPSRPSGAVSEDSILSESFDQLDLPERLPETPLRGCVSDNLRGLEQPPSEG---LK 596  
Db 534 ARASRPSGAVSEDSILSESFDQLDLPERLPETPLRGCVSDNLTGLEEPPSGGCLR 593  
Qy 597 RWMQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631  
Db 594 RWRQDPLGDSFSLTDCQEVTAAYRQALRVCSKLT 628

## RESULT 7

US-10-618-941-77  
; Sequence 77, Application US/10618941  
; Publication No. US20040197792A1  
; GENERAL INFORMATION:  
; APPLICANT: WHYTE, DAVID  
; APPLICANT: MANNING, GERARD  
; APPLICANT: CAENEPEEL, SEAN

; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 77
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-618-941-77

Query Match 85.8%; Score 2824.5; DB 17; Length 672;
Best Local Similarity 85.7%; Pred. No. 5.9e-176;
Matches 544; Conservative 26; Mismatches 54; Indels 11; Gaps 3;

Qy 1 MESVALLQRPQAPSAASALASESARPLADGLIKSPKPLMKQAVKRRHHKHLRHYEFL 60
Db 45 MESLVFARRSGPTPS---AAELARPLAELGKSPKPLMKQAVKRRHHKHLRHYEFL 100

Qy 61 ETLGKGTGKVKKARESGRLVAIKSRKDKIQDEODLLHRRREIEMSSLNHPHIIAIIH 120
Db 101 ETLGKGTGKVKKARESGRLVAIKSRKDKIQDEODLLHRRREIEMSSLNHPHIIAIIH 160

Qy 121 EVFENSSKIVIMVEYASRGDLVDYISRRPLSRDARHFRQIVSALHYCHONGIVHRDL 180
Db 161 EVFENSSKIVIMVEYASRGDLVDYISRRQUSERARHFRQIVSAVHYCHONRVHRDL 220

Qy 181 KLENILLDANGNIKIADFGLSNLHYKGFLOTFGCGSPLYASPEIVNGKPYVGPVDSWSL 240
Db 221 KLENILLDANGNIKIADFGLSNLHYKGFLOTFGCGSPLYASPEIVNGKPYVGPVDSWSL 280

Qy 241 GVLLYILVHGTMPDQDHKTLVQISNGAYREPPKSDACGLIRWLLMNPTRATILED 300
Db 281 GVLLYILVHGTMPDGHDKTLVQISNGAYREPPKSDACGLIRWLLMNPTRATILED 340

Qy 301 VASHWWNWGTYTGVGOEALREGHPGSDGDFGRASMDWLRSSRPILLENGAKVCSFFKQ 360
Db 341 VASHWWNWGTYATRVGOEAPHEGGHFGSDSARASMDWLRSSRPILLENGAKVCSFFKQ 400

Qy 361 HVPGGSTVTPGLERQHSIKSRKENDMAQNIQGDPAEDTSSRPGKSSIKLPKILKKSS 420
Db 401 HAPGGSTVTPGLERQHSIKSRKENDMAQSLHSDTADTAHRPCKSNLKLPGKILKKVS 460

Qy 421 TSSEVEDPQELRPVPTPGQVPVAVSLLPKGIKKRQRESGYSSPEPSSEGLLD 480
Db 461 ASAEQVEDPPPELSPFASPQQAAP--LLPKGILKKRQRESGYSSPEPSSEGLLD 517

Qy 481 ASDVFVSGDPVEOKSPOASGLLHRKGLTLKNGKFSRTALEGTTPTFGSLDQLASSHPA 540
Db 518 AGDVVFSGDPKEQKPPQASGLLHRKGLTLKNGKFSOTALEAAPTTFGSLDELAAPRPL 577

Qy 541 ARPSRPGVSEDSILSESDQDLDELPERLPETPLRGCVSDNRLGLEQPPSEG---LK 596
Db 578 ARASRPGVSEDSILSESDQDLDELPERLPETPLRGCVSDNLTGLEPSPGSGCLR 637

Qy 597 RWMQESLGDSFSLTDCQEVTAARQALGICKSLK 631
Db 638 RWRQDPLGDSFSLTDCQEVTAATYRQALRVCSKLT 672

RESULT 8
US-10-311-034-12
; Sequence 12, Application US/10311034
; Publication No. US20040023242A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: BANDMAN, Olga
; APPLICANT: BOROWSKY, Mark L.

; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Yan
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: GREENWALD, Sara R.
; APPLICANT: RAMKUMAR, Javalaxmi
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: KEARNEY, Liam
; APPLICANT: BURFORD, Neil
; APPLICANT: NGUYEN, Dannel B.
; APPLICANT: TANG, Y. Tom
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: HE, Ann
; APPLICANT: THORNTON, Michael
; APPLICANT: HAPALIA, April
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: LO, Terence P.
; APPLICANT: KHAH, Farrah A.
; APPLICANT: RECIPON, Shirley A.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: DING, Li
; APPLICANT: GREYER, Megan
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: BATRA, Sajeev
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0125 PCT
; CURRENT APPLICATION NUMBER: US/10/311,034
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,073; 60/213,467; 60/215,651; 60/216,605; 60/218,372; 60/228,056
; PRIOR FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-0-0
; 25
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023242A1 4841542CD1
; US-10-311-034-12

Query Match 82.6%; Score 2720.5; DB 15; Length 594;
Best Local Similarity 87.3%; Pred. No. 3.1e-169;
Matches 521; Conservative 22; Mismatches 47; Indels 7; Gaps 2;

Qy 39 MKQAVKRRHHKHLRHYEFLETGKGTGKVKKARESGRLVAIKSRKDKIKDEQDL 98
Db 1 MKQAVKRRHHKHLRHYEFLETGKGTGKVKKARESGRLVAIKSRKDKIKDEQDL 60

Qy 99 LHIRREIEMSSLNHPHIIAIIHEVFENSSKIVIMVEYASRGDLVDYISRRPLSRDARH 158
Db 61 MHIRREIEMSSLNHPHIIAIIHEVFENSSKIVIMVEYASRGDLVDYISRRQLSEREARH 120

Qy 159 FFRQIVSALHYCHONGIVHRDLKLENILLDANGNIKIADFGLSNLHYKGFLOTFGCSPL 218
Db 121 FFRQIVSAVHYCHONRVHRDLKLENILLDANGNIKIADFGLSNLHYKGFLOTFGCSPL 180

Qy 219 YASPEIVNGKPYVGPVDSWSLGVLLYILVHGTMPDQDHKTLVQISNGAYREPPKPS 278
Db 181 YASPEIVNGKPYVGPVDSWSLGVLLYILVHGTMPDGHDKTLVQISNGAYREPPKPS 240

Qy 279 DACGLIRWLLMNPTRATILEDVASHWWNVNGYTTGVGEQALREGHPGSGDFGRASMD 338
Db 241 DACGLIRWLLMNPTRATILEDVASHWWNVNGYATRVGEQAPHEGGHFGSDSARASMD 300

339 WLRSSRLLENGAKVCSFFKQHPGGSTVPGLEHSLKSKRENDMAONLQDPAED 398  
301 WLRSSRLLENGAKVCSFFKQHPGGSTVPGLEHSLKSKRENDMAONLQDPAED 360  
399 TSSRPGSSSLKPKGILKKSSTSSGEVQEDPQELRPVDPFGPQPAVSLPRKGLIKK 458  
361 TAHRPGSNLKLPGILKKSASAGVQEDPPELSPASPQQAAP---LLPKGILK 417  
459 SRORESGYSSPEPSESGELLDADSVFVSGDPVQKQPOAGLLHRRKGLKNGKFSRT 518  
418 PRORESGYSSPEPSESGELLDAGVDFVSGDPKQKPPQASGLLHRRKGLKNGKFSQT 477  
519 ALEGTTPSTFGSLDQASSHPAAPSRSRGVSDSILSSSFQDLDLPERLPETPLRG 578  
478 ALELAAPTTFGSLDELAPPRLPARASRGVSDSILSSSFQDLDLPERLPETPLRG 537  
579 VSDNLRGLQPPSEG---LKRWQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631  
538 VSDNLTGLEPPEPSEGCGSCLRRWRQDPLGDSFSLTDCQEVTAAYRQALRVCSKLT 594  
RESULT 9  
US-10-322-281-26  
; Sequence 26, Application US/10322281  
; Publication No. US20040126762A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: 529452001000  
; CURRENT APPLICATION NUMBER: US/10/322,281  
; CURRENT FILING DATE: 2002-12-17  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 26  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-322-281-26

Query Match 82.4%; Score 2712.5; DB 16; Length 616;  
Best Local Similarity 84.3%; Pred. No. 1.1e-168;  
Matches 535; Conservative 25; Mismatches 52; Indels 23; Gaps 9;  
QY 1 MESVALLQRPQASASALASESRPLADGLIKSPKLMKKQAVKRRHHKHLNRHYEFL 60  
DB 1 MESLVFARRSGPTPS---AAELARPLAEGSLIKSPKLMKKQAVKRRHHKHLNRHYEFL 56  
QY 61 ETLGKTYGVKKARESGRLVAIKSRDKIKDEQDLHTRRIETIWMSSLNHPHIIAH 120  
DB 57 ETLGKTYGVKKARESGRLVAIKSRDKIKDEQDLHTRRIETIWMSSLNHPHIIAH 116  
QY 121 EVFENSKIVMYSAGRDLYDYSERPLSRDARHFRQIVSALHYCHQNGIVHRDL 180  
DB 117 EVFENSKIVMYSAGRDLYDYSIQ--QLSE--ARHFRQIVSAYHYCHQNVHRDL 172  
QY 181 KLENILDANIKIADPGLSNLHYKGFLOTFGCSPLIYASPEIVNGKPYVGPVDSWSL 240  
DB 173 KLENILDANIKIADPGLSNLHYKGFLOTFGCSPLY--DEIVNGKPYTGPVDSWSL 230  
QY 241 GVLLIYLVHGTMPDGDHDKTLVQISNGAYREPPKPSDACGLIRWLLMWNPTRRATLED 300  
DB 231 GVLLIYLVHGTMPDGDHDKTLVQISNGAYREPPKPSDACGLIRWLLMWNPTRRATLED 290  
QY 301 VASHWVNWVTTGVGBQEARREGHPSGDPGRASMDLWLRSSRPLLENGAKVCSFFKQ 360  
DB 291 VASHWVNWVYATRVGEQAF--GGHPGDSARASMDLWLRSSRPLLENGAKVCSFFKQ 348  
QY 361 HVPGGSGTVPGLERQHSLLKSKRENDMAONLQDPAEDTSRPGKSSILKPKGILKKS 420  
DB 349 HAPGGSGTTPGLERQHSLLKSKRENDMAONLQDPAEDTAHRPGKSNLKLPGILKKS 408

QY 421 TSSEVQEDPQELRPVDPFGPQPAVSLPRKGLKSKRQRESGYSSPEPSESGELLD 480  
DB 409 ASAGVQEDPPELSPDP---GQAAP---LLPKGILKPKRQRESGYSSPEPSESGELLD 463  
QY 481 ASDVFSVSGDPVEQKSPQASGLLHRRKGLKNGKFSRTALEGTTPTFGSLDQASSHPA 540  
DB 464 AGDVFSVSGDPKQKPPQASGLLHRRKGLKNGKFSQTALELAAPTTFGSLDELAPRPL 523  
QY 541 ARSRPSGAVSRSISLSSSFQDLDLPERLPETPLRGCVSDNLRGLQPPSEG---LK 596  
DB 524 ARASRPSGAVSRSISLSSSFQDLDLPERLPETPLRGCVSDNLTGLLEPPEPSEGCLR 583  
QY 597 RWOQESLGDSCFSLTDCQEVTAAYRQALGICSKLS 631  
DB 584 RWRQDPLGDSFSLTDCQEVTA--YQALRVCSKLT 616  
RESULT 10  
US-10-363-829-404  
; Sequence 404, Application US/10363829  
; Publication No. US20040142331A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, Stuart E.; Lincoln, Stephen E.;  
; APPLICANT: Altus, Christina M.; Dufour, Gerard E.;  
; APPLICANT: Chalup, Michael S.; Jackson, Jennifer L.;  
; APPLICANT: Jones, Anisea L.; Yu, Jimmy Y.;  
; APPLICANT: Wright, Rachel J.; Gietzen, Darryl;  
; APPLICANT: Liu, Tommy F.; Yap, Pierre E.;  
; APPLICANT: Dahl, Christopher R.; Momiya, Monika G.;  
; APPLICANT: Bradley, Diana L.; Rohatgi, Sameer D.;  
; APPLICANT: Harris, Bernard; Roseberry Lincoln, Ann M.;  
; APPLICANT: Gerstin, Jr., Edward H.; Peralta, Careyna H.;  
; APPLICANT: David, Marie H.; Panzer, Scott R.;  
; APPLICANT: Flores, Vincent Z.; Dafo, Abel;  
; APPLICANT: Marwaha, Rakesh; Chen, Alice J.;  
; APPLICANT: Chang, Simon C.; Au, Alan P.;  
; APPLICANT: Inman, Rebekah R.  
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT  
; FILE REFERENCE: PT-1183 USN  
; CURRENT APPLICATION NUMBER: US/10/363,829  
; PRIOR FILING DATE: 2003-03-05  
; PRIOR APPLICATION NUMBER: PCT/US01/27628  
; PRIOR FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: US 60/229,751  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: US 60/229,749  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: US 60/229,750  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: US 60/229,747  
; PRIOR APPLICATION NUMBER: US 60/229,748  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: US 60/230,583  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: US 60/230,517  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: US 60/230,610  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: US 60/230,597  
; PRIOR FILING DATE: 2000-09-06  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 506  
; SOFTWARE: PERL Program  
; SEQ ID NO 404  
; LENGTH: 406  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: LG:982800.1.orfl:2000SEP08  
US-10-363-829-404

Query Match 51.6%; Score 1699.5; DB 9; Length 661;

Best Local Similarity 55.5%; Pred. No. 1.7e-102;  
Matches 372; Conservative 74; Mismatches 151; Indels 73; Gaps 18;

Qy	14	PSAALASESARPLADGLIKSPKPMKQAVKRRHHKHNLRHRYEFLETGKGYGVKVK	73
Db	13	PDGLGAPGSPREAVAGATAALEP-RKPHGVKRRHHKHNLRHRYELOETLGKGYGVKVR	71
Qy	74	ARES-SGRVAIKSRDKIKDEODLHIREIEMSSLNPHPIIAIHEVFENSKIIV	132
Db	72	ATERFSGRVVAIKSRDKIKDEODMWHIREIEMSSLNPHPIISIVYFENKDKIVII	131
Qy	133	MEYASRGDLYIISERPRLSERDARHFRQIVSALHYCHONGIIVHRDLKENILLDANGN	192
Db	132	MEYASKGELYIISERRRLSERETRHFRQIVSAVHYCHONGIIVHRDLKENILLDANGN	191
Qy	193	IKIADFGSLNLYHKGFQTCGSPLYASPIVNGKPYVGVPEVDSWSLGLVLLYLVHGTM	252
Db	192	IKIADFGSLNLYQDKFQTCGSPLYASPIVNGRPRYRGPVDSWALGVLLYLVHGTM	251
Qy	253	PFQGDHKTIVKQISNGAYREPPKPSDACGLIRWLLMVNPTRRATLEDVASHWVNWGYT	312
Db	252	PFQGFHKNLIRQISSGEYREPTQPSDARGLIRWLLMVNPTRRATLEDIANHWWNWGYK	311
Qy	313	TGVGEQALREGGHPSCDFGRASMDWLRSSRPILLENGAKVCSFFKQHVPGGSGTVPGL	372
Db	312	SSVCDCCDALHDSPL---LARIIDWHHRSTGLQADTEAKMKGLAK---PTTSEVM--L	362
Qy	373	ERQSLKSKKENDMAQNLQDPAEDTSRPGKSSKLPKGLIKKKS-----STSSGEVQ	427
Db	363	ERQSLKSKKENDFAQSGQDAPES-----PSKLSKRKPKGLIKKRSNHRSHSTGFIE	418
Qy	428	-----EDPQEL-----RPVPTGQPVPAVS--LLPRKGLIKKRSQRESGY	466
Db	419	GVVGPALPSTFKMEQDLCTGTVLLPSSPEAEVPGKLSPKQATWPKGILKTKTQRESGY	478
Qy	467	YSSPEPSEGLLDASDVFSVGDVPEQKSP-----QASGLLLHRKGLIKLNGKFSRTAL	520
Db	479	YSSPERSESELDSNDVM--GSSITPSPDPARVTSLSLSCRRKGLIKHSSKYSAGTM	536
Qy	521	EGTTPS-----TFGSLDQ-----LASSHPAARSPRGSAVSEDSILSSSEFDDLDLPER	569
Db	537	DPALVSPEMPFTLSLSEPGVPAEGLRSY-----SRPSSVISDSDSVLSSDSFDDLDLQEN	591
Qy	570	LP-ETPLRGCVSDN-----LRGLEQPPSEGLKRWQESLGDSCFS-LTDCQEVTAAYR	621
Db	592	RPARQIRSCVSAENFLQIQDFEGLQNRPRPQYLKRYNRNLADSSFSLITDMDVTVQYK	651
Qy	622	QALGICSKLS 631	
Db	652	QALEICSKLN 661	

## RESULT 13

US-09-780-949-6  
; Sequence 6, Application US/09780949  
; Patent No. US20020006618A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapellier-Liberman, Rosana.  
; APPLICANT: Weich, Nadine S.  
; APPLICANT: Galvin, Katherine M.  
; TITLE OF INVENTION: Methods for Using 20893, a Human Protein  
; TITLE OF INVENTION: Kinase  
; FILE REFERENCE: 035800/209015  
; CURRENT APPLICATION NUMBER: US/09/780,949  
; CURRENT FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: US 60/181,690  
; PRIOR FILING DATE: 2000-02-09  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 661  
; TYPE: PRT  
; ORGANISM: H. sapiens

## US-09-780-949-6

Query Match 51.6%; Score 1699.5; DB 9; Length 661;  
Best Local Similarity 55.5%; Pred. No. 1.7e-102;  
Matches 372; Conservative 74; Mismatches 151; Indels 73; Gaps 18;

Qy	14	PSAALASESARPLADGLIKSPKPMKQAVKRRHHKHNLRHRYEFLETGKGYGVKVK	73
Db	13	PDGLGAPGSPREAVAGATAALEP-RKPHGVKRRHHKHNLRHRYELOETLGKGYGVKVR	71
Qy	74	ARES-SGRVAIKSRDKIKDEODLHIREIEMSSLNPHPIIAIHEVFENSKIIV	132
Db	72	ATERFSGRVVAIKSRDKIKDEODMWHIREIEMSSLNPHPIISIVYFENKDKIVII	131
Qy	133	MEYASRGDLYIISERPRLSERDARHFRQIVSALHYCHONGIIVHRDLKENILLDANGN	192
Db	132	MEYASKGELYIISERRRLSERETRHFRQIVSAVHYCHONGIIVHRDLKENILLDANGN	191
Qy	193	IKIADFGSLNLYHKGFQTCGSPLYASPIVNGKPYVGVPEVDSWSLGLVLLYLVHGTM	252
Db	192	IKIADFGSLNLYQDKFQTCGSPLYASPIVNGRPRYRGPVDSWALGVLLYLVHGTM	251
Qy	253	PFQGDHKTIVKQISNGAYREPPKPSDACGLIRWLLMVNPTRRATLEDVASHWVNWGYT	312
Db	252	PFQGFHKNLIRQISSGEYREPTQPSDARGLIRWLLMVNPTRRATLEDIANHWWNWGYK	311
Qy	313	TGVGEQALREGGHPSCDFGRASMDWLRSSRPILLENGAKVCSFFKQHVPGGSGTVPGL	372
Db	312	SSVCDCCDALHDSPL---LARIIDWHHRSTGLQADTEAKMKGLAK---PTTSEVM--L	362
Qy	373	ERQSLKSKKENDMAQNLQDPAEDTSRPGKSSKLPKGLIKKKS-----STSSGEVQ	427
Db	363	ERQSLKSKKENDFAQSGQDAPES-----PSKLSKRKPKGLIKKRSNHRSHSTGFIE	418
Qy	428	-----EDPQEL-----RPVPTGQPVPAVS--LLPRKGLIKKRSQRESGY	466
Db	419	GVVGPALPSTFKMEQDLCTGTVLLPSSPEAEVPGKLSPKQATWPKGILKTKTQRESGY	478
Qy	467	YSSPEPSEGLLDASDVFSVGDVPEQKSP-----QASGLLLHRKGLIKLNGKFSRTAL	520
Db	479	YSSPERSESELDSNDVM--GSSITPSPDPARVTSLSLSCRRKGLIKHSSKYSAGTM	536
Qy	521	EGTTPS-----TFGSLDQ-----LASSHPAARSPRGSAVSEDSILSSSEFDDLDLPER	569
Db	537	DPALVSPEMPFTLSLSEPGVPAEGLRSY-----SRPSSVISDSDSVLSSDSFDDLDLQEN	591
Qy	570	LP-ETPLRGCVSDN-----LRGLEQPPSEGLKRWQESLGDSCFS-LTDCQEVTAAYR	621
Db	592	RPARQIRSCVSAENFLQIQDFEGLQNRPRPQYLKRYNRNLADSSFSLITDMDVTVQYK	651
Qy	622	QALGICSKLS 631	
Db	652	QALEICSKLN 661	

## RESULT 14

US-10-354-358-82  
; Sequence 82, Application US/10354358  
; Publication No. US20030157082A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc  
; APPLICANT: Hunter, John Joseph  
; APPLICANT: MacBeth, Kyle J.  
; APPLICANT: Tsai, Fong-ying  
; APPLICANT: Lesoon, Andrea  
; APPLICANT: Lightcap, Eric S.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Rudolph-Owen, Laura A.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING  
; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,  
; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,  
; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,  
; TITLE OF INVENTION: 7522, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,



```

; TITLE OF INVENTION: 8990, 2100, 92888, 64598, 10480, 20893, 32330, 1586, 9943,
; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
; FILE REFERENCE: MPI02-02OP1NONNM11
; CURRENT APPLICATION NUMBER: US/10/354,358
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US 60/353,600
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/364,517
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/371,075
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/371,507
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/372,984
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/374,194
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/382,995
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/385,023
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/388,853
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/389,395
; PRIOR FILING DATE: 2002-06-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-354-358-82

```

Query Match	51.8%; Score 1699.5; DB 14; Length 661;
Best Local Similarity	55.5%; Pred. No. 1.7e+102;
Matches	372; Conservative 74; Mismatches 151; Indels 73; Gaps 18;
QY	14 PSASALASASARPLADGLIKSPKPLMKQAVKHEHHKHNLRHYEFLETLGKTYGKVKK 73
DB	13 PDGLGAGSPREAVAGATAALEP-RKPHGVKXHHKHNLRHYEQETLGKTYGKVKR 71
QY	74 ARES-SGRIVAIKSIKDKIKBDQDLHLHRRTEIMSSLNHPHIIAIEHFVFNSSKIVIV 132
DB	72 ATERFSGRVAIKSIKDKIKBDQDMHHIRREIEMSSLNHPHIIISIEVFENDKIVII 131
QY	133 MEYASRGDIYDISERPRISERDARHFFROIVSALHYCHONGIVHVRDLKLENILLDANGN 192
DB	132 MEYASKGELYDISERRRISERETRIHFFROIVSAMVYCHNGVVRDLKLENILLDDNCN 191
QY	193 IKIADFGLSNLYHKGFLOTFCGSPLYASPEIYNGRPYGVPEVDSWGLVLLYILVHGTW 252
DB	192 IKIADFGLSNLYQDKFLOTFCGSPLYASPEIYNGRPYGVPEVDSWALGVLLYLVYGTW 251
QY	253 PDGQDHKTLVKQISNGAVREPPKPSDACGLIHWLWVNPTRATIEDVASHWWNVNGYT 312
DB	252 PDGDFDHKNLIRQISSGEYREPTQPSDARGLISWMLMVPDRATTIEDIANHWVWVNGYK 311
QY	313 TGVGEQEARREGHPSGDFGRASWADWLRRSPPLLENGAKVCSFFKQHPVPGGISTVPLG 372
DB	312 SSVCDCCDALHDSSEPL-----LARIIDWHHRSTGLQADTEAKMKGLAK-----PTTSVM--L 362
QY	373 ERQHSLSKKRKENDMAQLQGDPAEDTSRPGKSSLLPKGILKKKS-----STSSGEVQ 427
DB	363 ERQHSLSKKKENDPQASGQDAVPES-----PSKLSKKRPKGIILKKRSNHSRHSHTGFIEA18 418
QY	428 -----EDPQEL-----RPVPDPTGPQVPAPVS-LLPKRGILKKSRQRESGY 466
DB	419 GYVGPALPSTFKMEQDLCRTGVLLPSSPEAEVPGKLSPKQASATMPKGIILKTKQRESGY 478
QY	467 YSSPEPSESGELLDASDVVFVSGDPVEQKSP-----QASGLLHHRKGIILKLNCKGFSRTAL 520

```

479 YSSPSESSELDSNDVM--GSSISPSPGPPDPARVTSHSLCRRKIGILKHSKYSAGTM 536
QY 521 EGTTPS-----TFGSLDQ-----LASSHPAARSPRGAVSDESDLSSESFDQLDLPER 569
Db 537 DPALVSEMPFLESLESPGVPAAGLSRSY-----SRPSSVIGDDSVLSSDSFDLLDLOEN 591
QY 570 LP-ETPLRGCVSDN-----LRLGQPPEGLKRWQESLGDSCFS-LTDCQEVTAAYR 621
Db 592 RPARQRIRSCVSAENFLIQDFEGLQNRPRPQYLKRYNRLADSPSLTDDMDVTVQVK 651
QY 622 QALGICSKLS 631
Db 652 QALEICSKLN 661

RESULT 15
US-10-723-860-2553
; Sequence 2553, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Composition
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2553
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-2553

```

Query Match	51.6%;	Score 1699.5;	DB 17;	Length 661;
Best Local Similarity	55.5%;	Pred. No. 1.7e-102;		
Matches	372;	Conservative	74;	Mismatches 151; Indels 73; Gaps 18;
Qy	14	PSASALASASARPLADGLIKSPKPLMKQAVRRHHHKNLRHRYEPLETGLKGTGYCKVKK	73	
Db	13	PDLGAGPGSPREAVAGATAALEP-RKPHGVRRHHHKNLRHRYELQETGLKGTGYCKVKR	71	
Qy	74	ARES-SGRVLAKTSRKOKIKDEQDLHTRRRIETMSSLNHHPHIATIHVEFVENSXKIVV	132	
Db	72	ATERFSGRVLAKTSRKOKIKDEQDWHIRRIETMSSLNHHPHIISYIEVFENKDKIVII	131	
Qy	133	MEYASRGDLIDYISBRPLSRSDARHFRQIVSALHYCHQNGIVHRDLKENILLDANGN	192	
Db	132	MEYASKGELIDYISERRLSERETHFRQIVSAVHYCHKNGVHRDLKENILLDDNCN	191	
Qy	193	IKIADFGLSNLVHKKFTQTCGSPLYIASPETVNGKPYVGPVEDVSNLSGLLVLYLVHGTW	252	
Db	192	IKIADFGLSNLVKKQKFLQTCGSPLYIASPEIVNGRPYRGPEVDSWALGVLVLYLVGTW	251	
Qy	253	PDGQDHKTIVQISNGAYREPKSDACGLIRWLLMVNPTRRATLEDVASHWWNVNGYT	312	
Db	252	PDGDFDHKNLIRQISSGEYREPTQSDARGLLIRWLLMVNPNDRRATIEDIANHWWNVNGYK	311	
Qy	313	TGVGQEOALREGCHPSGDGFGRAMADWLRSSRPILLENGAKVCSPFKQHVPVGGSTVPGL	372	
Db	312	SSVCDCDALHDSSESPL----LARIIDWHHRSTGLQADTEAKWKGLAK---PITSEVM--L	362	
Qy	373	ERQSHLKKRSKENDMAQNQLQGDPAEDTSSRPGKSSKLKPKGILKKKS-----STSSGEVQ	427	
Db	363	ERQRSILKKSKKENDFAQSQGDVAPES---PSKLSKRKPKGILKKRSNSEHRSHSTGFIE	418	
Qy	428	-----BDQDEL-----RVPDTPGQVPVAVS--LLPRKGILKKRSQRRESGY	466	
Db	419	GVVGGPALPSTFMQEODLCRTGVLLPSPSPAEBYVGKLSPKQSATMPKPKGILKKTQQRRESY	478	

```

Qy 467 YSSPESGELLASDVFSQDPVEQKSP-----QASGLLLHRKGIKUNGKFSRTAL 520
Db 479 YSSPESGELLASDVFSQDPVEQKSP-----QASGLLLHRKGIKUNGKFSRTAL 520
Qy 521 EGTTPS---TFGSLDQ-----LASSHPAARPSRPSGAVSEDSILSSSEFDOLDLPER 569
Db 537 DPALVSPMPPTLESUSEPCVPAEGLRSY-----SRPSSVSDSDSVLSSDSFDLLDIQEN 591
Qy 570 LP-ETPLRGCVSDN-----LRGLEQPPSEGLKRWQESLGDSGFS--LTDCEVTAAYR 621
Db 592 RPARQIRSCVSAENFLQIQDFEGLQNRPRPQYLKRYENRLADSSFSLLTMDDDVTQVYK 651
Qy 622 QALGICSKLS 631
Db 652 QALEICSKLN 661

```

Search completed: January 22, 2005, 07:00:33  
Job time : 150.329 secs